

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة 8 ماي 1945 قالمة
Université 8 Mai 1945 Guelma

Faculté des Sciences de la Nature et de la Vie, Sciences de la Terre et de l'Univers



Mémoire En Vue de l'Obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité/Option : Biochimie Appliquée

Département : Biologie

**Thème : Analyse des gènes différentiellement exprimés dans le cancer
des poumons basée sur des outils bioinformatiques.**

Présenté par : Mlle Aidi Hadil

Mlle Amiri Messaouda

Mlle Belhassab Norelhouda.

Devant le jury composé de :

Présidente : Mme BOUSSENANE H.	MCB	Université de Guelma
Examineur : Mr MOKHTARI AH.	MCB	Université de Guelma
Encadreur : Mr OUMEDDOUR A.	MCA	Université de Guelma

Juin 2023

Remerciements

*On tient tout d'abord à remercier le bon Dieu pour nous avoir donné la santé,
le courage et la volonté pour réaliser ce travail.*

*Nous remercions Madame «Boussenane H », Maître de
conférences à l'Université 8 mai 1945 Guelma, qui nous fait l'honneur de
présider le jury et de juger notre travail.*

*Nous remercions également Monsieur «Mokhtari Abdelhamid», maître de
conférences à l'Université 8 mai 1945 Guelma, d'avoir accepté de faire partie
du jury et d'évaluer notre travail.*

*On tient à exprimer notre profonde gratitude et nos sincères remerciements
à notre professeur encadrant monsieur «Oumeddour Abdelkader», pour toute
l'aide qu'il nous a apporté en suivant notre travail avec sérieux et intérêt, et
pour sa patience, sa générosité et ses conseils.*

*Nous souhaitons également exprimer nos sincères remerciements à tous nos
enseignants pour tout ce que nous avons appris avec eux tout au long de
notre parcours universitaire.*

Dédicaces

Tout d'abord ,je tient à remercier **DIEU** de m'avoir donné le courage et la volonté de terminer ce mémoire

Chère maman, Aujourd'hui, au moment de conclure ce mémoire de fin d'études, je me tourne vers toi avec une gratitude infinie pour tout ce que tu as fait pour moi. Tu as été bien plus qu'une mère pour moi, tu as également joué le rôle d'un père . Tu as été présente à chaque étape de ma vie, soutenant mes rêves, je t'aime

Cher papa, malgré ton absence physique, ton esprit demeure vivant dans mon cœur. Je continuerai à avancer . Tu me manques énormément, et je t'aimerai toujours. Repose en paix papa. Cette réussite est également la tienne.

*À mon frère bien-aimé **Abd enour**, ainsi qu'à mes sœurs **Loubna** et **Asma**, Votre présence et votre soutien tout au long de ma vie ont été inestimables. Que Dieu vous accorde le succès dans votre vie*

À mes précieuses copines, 'Hadil 'et 'Hana' (messaouda), nous avons partagé des moments de travail acharné, je souhaite que notre amitié continue de s'épanouir au-delà de nos études, que nous puissions partager de nouveaux défis et célébrer ensemble de nombreux succès à venir. merci d'avoir rendu cette expérience si spéciale.

Enfin, je voudrais remercier tous ceux qui ont apporté leur aide et leur soutien tout au long de mes études.

À vous tous, je vous dédie cette réalisation avec une profonde gratitude et reconnaissance. Votre impact sur ma vie restera gravé dans ma mémoire pour toujours.

NOR ELHOUDA

Avec l'aide et la protection d'ALLAH et d'un cœur plein d'amour,

Je dédie mon travail :

*À ma perle précieuse ,ma fierté ,la femme de fer "Ma Mère" **Batah Nassima**
qui a joué le rôle d'un père et d'une mère à la fois malgré toutes les difficultés
,elle m'a encouragé, aidé et guidé dans le chemin de la vie. Je remercie
infiniment le Tout Puissant ALLAH de m'avoir donné une mère comme toi.*

Que Dieu te bénisse et te protège, je t'aime trop.

*À l'esprit pur de mon cher oncle "**Batah el Djayed** " , qui était comme un père pour moi ,
que Dieu le protège dans Son vaste paradis .*

*À ma grande sœur "**Amiri Amira**" qui n'a pas cessé de me conseiller,
encourager et soutenir tout au long de mes études. Merci pour tes efforts, que
Dieu te bénisse ma douce sœur.*

*À mon adorable petite sœur "**Amiri achouak** " ma complice, ma meilleure
amie, qui sait toujours comment me procurer la joie et le bonheur. Tu es une
source infinie de joie, de tendresse. Je te souhaite beaucoup de joie et de réussite
à ton baccalauréat.*

*À mes chers neveux **ghiles** et **razane**, mes anges, vous êtes une belle source de
joie et de détresse. Je vous aime.*

*À mes chères copines "**Hadil** " et "**Nor el houda** " , je vous dis merci du fond du cœur
pour votre amitié sincère, ainsi que pour tous les beaux moments et les efforts donnés
pour avoir un très bon travail.*

*Et à chaque personne chère à mon cœur , qui m'a soutenu, de près ou de loin, merci
beaucoup.*

Amiri Messaouda

le parcours de 17 ans s'est achevé, avec ses moments mauvais et ses joies inoubliables. Ce chemin n'a ni été bref ni parsemé de facilités, mais je suis parvenue à le parcourir avec succès.

Tout d'abord, je tiens à exprimer ma fierté et ma profonde gratitude envers moi-même pour mes efforts et ma patience tout au long de ce parcours. Et avec une gratitude sincère, je dédie cette réalisation à :

Mon cher père, qui m'a énormément aidé à arriver là où je suis aujourd'hui. Merci infiniment pour tes sacrifices, et tout ce que tu as fait pour moi.

Ma mère, ma source d'inspiration constante, mon pilier de force et ma plus grande supportrice. Il n'y a pas de mots assez forts pour te remercier de tout ce que tu m'as apporté en termes d'encouragement et d'aide. Ton dévouement et ta générosité ne seront jamais oubliés. Je t'aime plus que les mots ne pourraient jamais l'exprimer.

Mes frères bien-aimés, je suis profondément reconnaissante d'avoir des frères merveilleux comme vous. Que Dieu vous bénisse abondamment dans vos études et dans votre vie.

Je dédie ce travail à la mémoire de **ma chère grand-mère maternelle**. Que Dieu préserve son âme dans Son vaste paradis.

À toute ma famille, et en particulier à mes chères cousines maternelles, ainsi qu'à toutes mes amies et à tous mes proches.

À mes chères copines Nor Elhouda et Hana, je tiens à vous exprimer ma gratitude pour tous les bons moments que nous avons partagés tout au long de ce parcours. Vos efforts, votre travail acharné et votre aide sont été inestimables.

Et à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail, je vous dis merci et je vous suis sincèrement reconnaissante.

Hadil

Résumé

Le cancer des poumons est une maladie caractérisée par la croissance incontrôlée de cellules anormales dans les tissus pulmonaires. Le traitement du cancer des poumons implique généralement une combinaison de thérapies telles que la chirurgie, la radiothérapie, la chimiothérapie et la thérapie ciblée, en fonction du stade et du type de cancer. Le but de notre étude vise à identifier les gènes clés communs liés au cancer des poumons et à son diagnostic. Trois profils de données ont été téléchargés depuis la base de données GEO, puis ils ont été analysés à l'aide de plusieurs outils bioinformatiques. Les gènes à expression différentielle commune (DEG) ont été déterminés par l'outil Venn Diagram. L'analyse de *Gene Ontology* (GO) et l'analyse de *Kyoto Encyclopedia of Genes and Genomes* (KEGG) ont été effectuées par site DAVID. L'interaction protéine-protéine (PPI) a été créée à l'aide du site STRING et visualisée par le logiciel Cytoscape. La survie globale a été examinée par l'outil Kaplan-Meier Plotter, enfin l'expression des gènes clés a été vérifiée à l'aide de la base de données GEPIA2.

Nous avons identifié 179 DEG, comprenant 78 gènes surexprimés et 101 gènes sous-exprimés. Les résultats de l'analyse GO et KEGG ont révélé une forte association de ces gènes avec plusieurs processus biologiques, cellulaires et moléculaires liés aux cancer des poumons. En outre, nous avons repéré huit gènes clés exprimés de manière significative dans les tissus cancéreux par rapport aux tissus normaux, qui ont été associés à une mauvaise survie globale des patients. En résumé, ces huit gènes clés (ASPM, BUB1, KIF11, CCNA2, AURKB, NDC80, CCNB1 et CCNB2) peuvent constituer des bons biomarqueurs pour le diagnostic et la thérapie ciblée du cancer des poumons.

Mots clés : cancer des poumons, bioinformatique, gènes à expression différentielle, diagnostic, thérapie ciblée.

Abstract

Lung cancer is a disease characterized by the uncontrolled growth of abnormal cells in lung tissues. The treatment of lung cancer typically involves a combination of therapies such as surgery, radiation therapy, chemotherapy, and targeted therapy, depending on the stage and type of cancer. The aim of our study was to identify common key genes associated with lung cancer and its diagnosis. Three data profiles were downloaded from the GEO database and subsequently analyzed by various bioinformatics tools. The genes with common differential expression (DEG) were determined by the Venn Diagram tool. The Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were performed using the DAVID website. The protein-protein interaction (PPI) network was created with the STRING website and visualized with the Cytoscape software. Overall survival was examined by the Kaplan-Meier Plotter. Finally, the expression of key genes was checked using the GEPIA2 database.

We have identified 179 DEG, including 78 upregulated genes and 101 downregulated genes. The results of the GO and KEGG analyses have revealed a strong association of these genes with several biological, cellular, and molecular processes related to lung cancer. In addition, we identified eight key genes expressed significantly in cancer tissue compared to normal tissues, which have been associated with poor overall patient survival. In summary, these eight key genes (ASPM, BUB1, KIF11, CCNA2, AURKB, NDC80, CCNB1 and CCNB2) may constitute good biomarkers for the diagnosis and targeted gene therapy of lung cancer.

Keywords : lung cancer, bioinformatics, differential expression genes, diagnosis, targeted therapy.

ملخص

سرطان الرئة هو مرض يتميز بنمو خلايا غير طبيعية في أنسجة الرئتين بشكل غير مسيطر عليه. يشمل علاج سرطان الرئة عادة مجموعة من العلاجات مثل الجراحة، العلاج الإشعاعي، العلاج الكيميائي والعلاج المستهدف وذلك وفقا لمرحلة ونوع السرطان. الهدف من دراستنا كان تحديد الجينات المشتركة الرئيسية المرتبطة بسرطان الرئة وتشخيصه. تم تنزيل ثلاثة ملفات بيانات من قاعدة بيانات GEO ثم تم تحليلها باستخدام العديد من أدوات المعلوماتية الحيوية. تم تحديد الجينات ذات التعبير الفرقي المشترك (DEGs) بواسطة قاعدة بيانات Venn Diagram. تم إجراء تحليل Gene Ontology (GO) وتحليل Kyoto Encyclopedia of Genes and Genomes (KEGG) باستخدام موقع DAVID. تفاعل بروتين-بروتين (PPI) تم إنشاؤه بواسطة موقع STRING و تم تصويره باستخدام برنامج Cytoscape. تم دراسة البقاء الشامل باستخدام أداة Kaplan Meier Plotter. وأخيرا تم التحقق من تعبير الجينات الرئيسية باستخدام قاعدة بيانات GEPIA2.

حددنا 179 DEGs، بما في ذلك 78 جينة مفرطة التعبير و 101 جينة غير مفرطة التعبير، كشفت نتائج تحليل GEO و KEGG عن ارتباط قوي لهذه الجينات بالعديد من العمليات البيولوجية، الخلوية و الجزيئية المرتبطة بسرطان الرئة. بالإضافة إلى ذلك، حددنا ثمانية جينات رئيسية تم التعبير عنها بشكل كبير في الأنسجة السرطانية مقارنة بالأنسجة الطبيعية، و التي ارتبطت بضعف بقاء المريض على قيد الحياة. باختصار، يمكن أن تشكل هذه الجينات الثمانية الرئيسية (ASPM, BUB1, KIF11, CCNA2, AURKB, NDC80, CCNB1, CCNB2) مؤشرات حيوية جيدة للتشخيص و العلاج المستهدف لسرطان الرئة.

الكلمات المفتاحية: سرطان الرئة، المعلوماتية الحيوية، الجينات ذات التعبير التفاضلي، التشخيص، العلاج المستهدف.

Table des matières

Remerciements

Dédicaces

Résumé

Abstract

ملخص

Liste des Figures

Liste des Tableaux

Liste des abréviations

Introduction 15

Étude bibliographique

1. Cancer du poumon 2

1.1. Anatomie des poumons 2

1.2. Définition 3

1.3. Epidémiologie des cancers du poumon 3

1.3.1. Dans le monde 3

1.3.2. Dans l'Algérie 3

1.4. Les types de cancer du poumon 4

1.4.1. Le cancer bronchique à petites cellules 4

1.4.2. Les cancers bronchiques non à petites cellules 4

1.4.1.2. Les Carcinomes épidermoïdes 5

1.4.1.3. Les carcinomes à grandes cellules 6

1.5. Les Symptômes 6

1.5.1. Symptômes liées au métastases 7

1.5.2. Syndrome paranéoplasique 7

1.6. Les Facteurs de risque 7

1.6.1. Le tabagisme 7

1.6.2. Le radon 8

1.6.3. les expositions professionnelles 8

1.6.4. pollution atmosphérique 8

1.6.5. Les antécédents personnels et familiaux 8

1.7. Les stades de cancer du poumon 9

1.7.1. La classification TNM 9

1.8. Diagnostic et traitement de cancer du poumon 12

1.8.1. Diagnostic 12

1.8.1.1. Examen clinique 12

1.8.1.2. Examen complémentaire 12

1.8.2. Traitement 13

1.8.2.1. La chirurgie 13

1.8.2.2. La chimiothérapie 13

1.8.2.3. La radiothérapie 13

1.8.2.4. La thérapie ciblée : 13

1.9. Prévention	14
<u>Etude bioinformatique</u>	
2. Méthodes	15
2.1. L'objectif de travail	15
2.2. Méthode d'analyse	15
2.2.1. Source des données et informations sur les puces à ADN	15
2.2.2. Traitement des données et identification des DEG	15
2.2.3. Étude de l'ontologie des gènes et de l'enrichissement des voies	15
2.2.4. Utilisation de la PPI pour l'analyse des réseaux et des modules	16
2.2.5. Sélection des gènes clés	16
2.2.6. Analyse de survie globale et validation de l'expression des gènes clés	16
3. Résultats et discussion	17
3.1. Résultats	17
3.1.2. Étude de l'ontologie génique et de l'enrichissement des voies ..	19
3.1.3. Étude de Kyoto Encyclopedia of Genes and Genomes (KEGG)	22
3.1.4. Analyse de l'interaction protéine-protéine	23
3.1.5. Sélection des gènes clés	24
3.1.6. Analyse de survie globale et validation de l'expression des gènes clés	25
3.2. Discussion	27
<u>Conclusion</u>	31
<u>Glossaire</u>	32
<u>Références bibliographiques</u>	34

Liste des figures

Figure 01 : Anatomie des poumons	2
Figure 02 : (A) Le nombre de nouveau cas de cancer bronchique en 2020, femmes et hommes. (B) le nombre de nouveau cas en 2020, hommes.	4
Figure 03 : Vue microscopique de l'adénocarcinome.	5
Figure 04 : Vue microscopique de carcinome épidermoïde.	5
Figure 05 : Vue microscopique de carcinomes à grande cellule	6
Figure 06 : Poumon d'un fumeur et d'un non fumeur.	7
Figure 07 : Les tracés volcaniques (<i>Volcano plots</i>) des trois ensembles de données GEO, le bleu représente les gènes sous-exprimés, le rouge représente les gènes surexprimés ($P < 0:05$ et $\log FC > 1.5$).	17
Figure 08 : Les DEG communs aux trois <i>datasets</i> . (A) les 78 gènes surexprimés. (B) les 101 gènes sous-exprimés.	18
Figure 09 : Le réseau PPI, les gènes surexprimés en jaune, les gènes sous-exprimés en vert.	23
Figure 10 : Les huit gènes clés classés par les algorithmes topologiques	24
Figure 11 : Analyse de la survie des huit gènes exprimés dans le cancer des poumons. Les tissus cancéreux en rouge, les tissus normaux en gris	25
Figure 12 : La vérification de l'expression des gènes clés dans les tissus de cancer des poumons par l'outil GEPIA2, (LUSC : <i>lung squamous cell carcinoma</i>).	26

Liste des tableaux

Tableau 01 : Répartition par stade selon la 8 ^{ème} classification TNM pour les cancers pulmonaires	9
Tableau 02 : Proposition des descripteurs T,N et M pour la huitième édition de la classification TNM pour le cancer du poumon.	10
Tableau 03 : Liste des DEG communs aux trois <i>datasets</i>	19
Tableau 04 : Analyse GO des DEG communs dans le cancer du poumon.	20
Tableau 05 : Analyse KEGG des DEG courants dans le cancer du poumon.	22

Liste des abréviations

AD :	Adénocarcinome.
ADN :	Acide Désoxyribonucléique.
ALK :	<i>Anaplastic lymphoma kinase.</i>
ARNm :	Acide ribonucléique messenger.
ASPM :	<i>Abnormal Spindle Microcephaly Associated.</i>
ATP :	<i>Adénosine Triphosphate.</i>
AURKB :	<i>Aurora kinase B.</i>
BP :	<i>Biological Process.</i>
BRAF1 :	<i>B-Raf Proto-Oncogene, Serine/Threonine Kinase.</i>
BUB1 :	<i>BUB1 mitotic checkpoint.</i>
CBP :	Cancer broncho-pulmonaire.
CBPC :	Le cancer bronchique à petites cellules.
CBNPC :	Cancer bronchique non à petites cellules.
CC :	<i>Cellular Component.</i>
CCNA2 :	<i>Cyclin A2.</i>
CCNB1 :	<i>Cyclin B1.</i>
CCNB2 :	<i>Cyclin B2.</i>
CIRC :	Centre international de recherche contre le cancer.
DAVID :	<i>Database for Annotation Visualization and Integrated Discovery.</i>
DEG :	<i>Differentially Expressed Genes.</i>
EGFR :	Récepteur du facteur de croissance épidermique.
FC :	<i>Fold change.</i>
GEO :	<i>Gene Expression Omnibus.</i>
GEPIA2 :	<i>Gene Expression Profiling.</i>
GO :	<i>Gene Ontology.</i>
GPL :	<i>Gene Platform.</i>
GSE :	<i>Gene Sample Expression</i>
HG :	<i>Humain Génome.</i>
ICP :	Irradiation crânienne prophylactique.
IL10 :	Interleukine 10.
IRM :	Imagerie par résonance magnétique.
KEGG :	<i>Kyoto Encyclopedia of Genes.</i>
KIF11 :	<i>Kinesin family member 1.1</i>
Log :	Logarithme.
MCC :	<i>Maximal Clinique Centrality.</i>
MF :	<i>Molecular Function.</i>
NDC80 :	<i>kinetochore complex component</i>
OMS :	Organisation mondiale de la santé .
P :	Valeur de probabilité.
PPI :	<i>Protein-Protein Interaction.</i>

SIADH :	Syndrome de sécrétion inappropriée d'hormone antidiurétique.
SPN :	Les syndromes paranéoplasiques.
TCA :	Temps de céphaline activée.
TDM :	Tomodensitométrie.
TEP :	La tomographie par émission de positrons.
TNM :	Tumeur, Ganglion Lymphatique, Métastase.
TP :	Taux de prothrombine.
UICC :	Union International Contre le cancer
VEGF :	<i>Vascular Endothelial Growth Factor</i> .

Introduction

Introduction

Le cancer du poumon est la tumeur maligne la plus courante et la principale cause de décès liée au cancer dans le monde entier [1]. Le tabac est le facteur de risque principal du cancer bronchique primitif (CBP), les autres facteurs de risque sont secondaires mais donnent également le cancer du poumon tels que les expositions professionnelles, la pollution atmosphérique et les antécédents familiaux [2]. On classe les cancers du poumon en deux grandes catégories : les cancers bronchiques non à petites cellules (CBNPC) (adénocarcinome, carcinome épidermoïde, carcinome à grandes cellules, etc.) et les cancers bronchiques à petites cellules (CBPC) [3]. Le traitement du cancer broncho-pulmonaire repose sur la chirurgie, la radiothérapie, la chimiothérapie associée ou non à la thérapie ciblée [4].

Les études d'expression génique sont un outil important pour l'analyse transcriptomique d'un organisme qui aide à quantifier les niveaux d'expression des gènes dans des conditions normales et pathologiques. Les profils d'expression génique de deux conditions différentes (*pathologique versus normal*) peuvent être comparés pour révéler des régulateurs clés potentiels ou des gènes différentiellement exprimés (DEG) [5]. L'analyse de l'expression génique a montré de nouveaux biomarqueurs potentiels et des cibles précieuses pour la gestion et le traitement des patients [6].

Notre étude repose sur l'utilisation de plusieurs outils bioinformatiques pour extraire et analyser un ensemble de données et identifier les DEG associés au cancer bronchique. Nous avons également effectué une analyse d'enrichissement des gènes (*Gene Ontology : GO*) et une analyse de la *Kyoto Gene and Genomic Encyclopedia (KEGG)* pour étudier leur implication dans le développement de cette maladie. Nous avons construit un réseau d'interaction protéine-protéine (*protein-protein interaction : PPI*) à l'aide de STRING et nous l'avons visualisé par Cytoscape, puis nous avons utilisé le plug-in Cytohubba pour identifier les gènes clés associés au cancer bronchique. Enfin, nous avons vérifié l'expression de ces gènes chez les patients atteints de cette maladie et leur impact sur les chances de survie globale (*Overall Survival : OS*) à l'aide de la base de données GEPIA2.

Ce mémoire comporte deux parties :

- ✓ Une partie théorique concerne le cancer bronchique : structure des poumons, épidémiologie, classification, facteurs de risque, les stades, traitement et diagnostic.
- ✓ Une partie bioinformatique concentrée sur l'analyse par plusieurs bases de données et logiciels des données des patients atteints de cancer bronchique.

Étude bibliographique

1. Cancer du poumon

1.1. Anatomie des poumons

Les deux poumons sont des organes de forme conique situés dans la cavité thoracique. Ils sont séparés par le cœur et d'autres structures du médiastin. Deux feuillets de séreuse qui forment la plèvre protègent chaque poumon. Chaque poumon est divisé en lobes pulmonaires par des profondes rainures appelées scissures (Figure 1). La scissure oblique du poumon gauche sépare le lobe supérieur du lobe inférieur. Dans le poumon droit, la scissure oblique et la scissure horizontale séparent le lobe supérieur, le lobe moyen et le lobe inférieur [7]. Ils sont reliés à la trachée par deux bronches souches centrales qui se séparent dans chacun des lobes en bronches puis en bronchioles. Les bronchioles se terminent par des alvéoles pulmonaires dans lesquelles des capillaires permettent les échanges gazeux d'oxygène (O_2) et de dioxyde de carbone (CO_2) entre l'air et le sang [8].

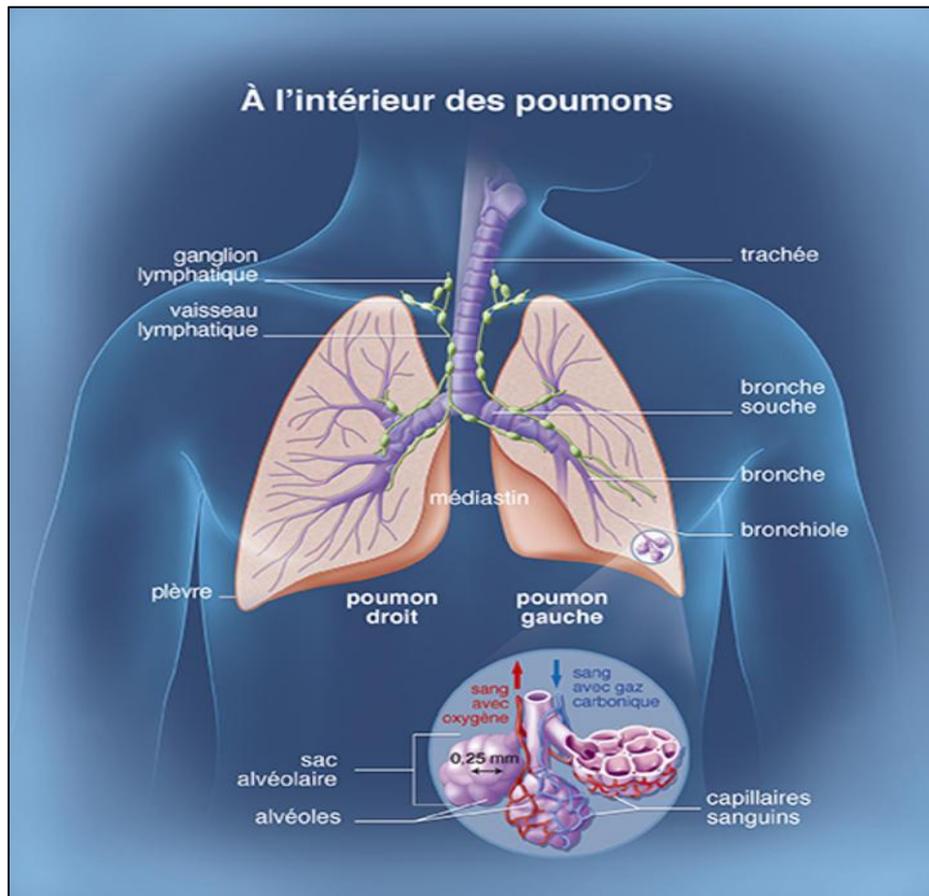


Figure 01: Anatomie des poumons [9].

1.2. Définition

Le cancer du poumon également connu sous le nom de cancer bronchique ou cancer broncho-pulmonaire, est une maladie des cellules des bronches ou plus rarement des cellules qui tapissent les alvéoles pulmonaires [9]. Il résulte de la prolifération incontrôlée d'une seule cellule qui en se multipliant et forme éventuellement une tumeur qui n'envahit initialement que l'organe d'où elle a pris naissance. Puis la tumeur peut s'étendre aux organes avoisinants. Elle peut aussi former des métastases, c'est-à-dire se propager aux organes à distance et causer la mort [10].

1.3. Epidémiologie des cancers du poumon

1.3.1. Dans le monde

Le cancer du poumon est le cancer le plus fréquent dans le monde [11]. Il constitue actuellement l'un des principaux problèmes de santé publique, avec une incidence mondiale de 1,2 million de nouveaux cas par an et une mortalité touchant 1,3 million de personnes [12]. En 2012, environ 1,6 million de personnes sont décédées des suites d'un cancer du poumon dans le monde (63 % chez les hommes, 57 % chez les femmes) [13]. Le cancer du poumon a l'incidence la plus élevée dans les pays en développement où le tabagisme est le plus répandu, est le plus courant dans 37 pays, dont la Russie, la Chine et une grande partie de l'Europe de l'Est, du Moyen-Orient et de l'Asie du Sud-Est [14].

1.3.2. Dans l'Algérie

En 2017, l'Algérie a enregistré 43 920 nouveaux cas de cancer, soit une incidence de 112,2 nouveaux cas pour 100 000 habitants. Le cancer colorectal est le cancer le plus fréquent chez les hommes, suivi du cancer du poumon (13,5 nouveaux cas pour 100 000 hommes) [15]. En 2020, l'Algérie a enregistré 58 418 cas de cancer, avec environ 4 774 nouveaux cas de cancer du poumon. Il occupe la première position chez les hommes avec 4 159 cas et 4 333 décès. Chez la femme, le cancer du poumon est rare (Figure 2) [16].

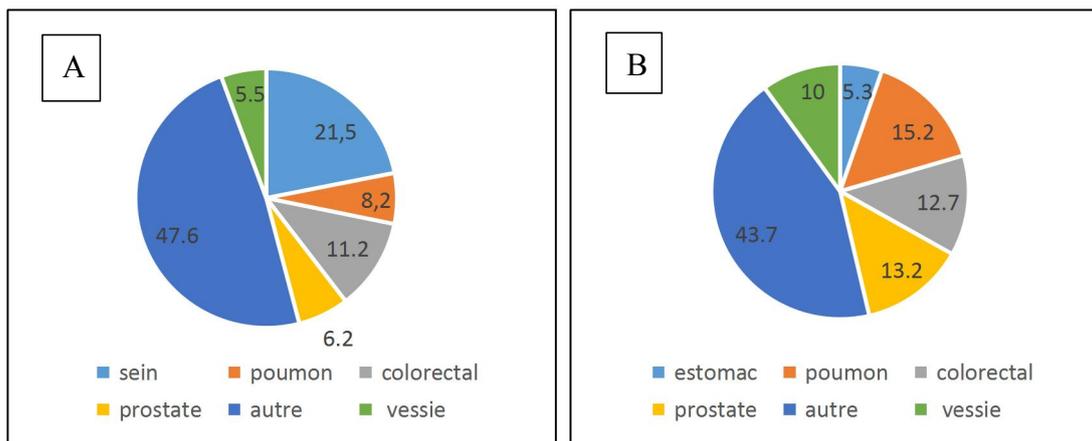


Figure 02: (A) Le nombre de nouveau cas de cancer bronchique en 2020, femmes et hommes. (B) le nombre de nouveau cas en 2020, hommes [16].

1.4. Les types de cancer du poumon

Les deux principaux types histologiques de CBP sont [17] :

1.4.1. Le cancer bronchique à petites cellules

Le cancer bronchique à petites cellules (CBPC) représente environ 15 % de tous les cancers du poumon et se caractérise par un taux de prolifération exceptionnellement élevé [18]. Le CBPC est caractérisé par un temps de doublement rapide et des métastases précoces et généralisées [19]. Il est généralement divisé en deux étapes, limitées et étendues [20].

1.4.2. Les cancers bronchiques non à petites cellules

Les cancers bronchiques non à petites cellules (CBNPC) représentent 85% des cancers bronchiques. Les types histologiques prédominants sont les adénocarcinomes, les carcinomes épidermoïdes et les carcinomes à grandes cellules [21].

1.4.2.1. Les adénocarcinomes

Les adénocarcinomes sont des tumeurs épithéliales présentant une différenciation glandulaire ou une sécrétion de mucines par les cellules tumorales [22]. Ils sont la classe majoritaire des CBNPC avec 40 à 50%. Environ 15% des cas sont associés à une mutation de l'EGFR (Figure 3) [23].

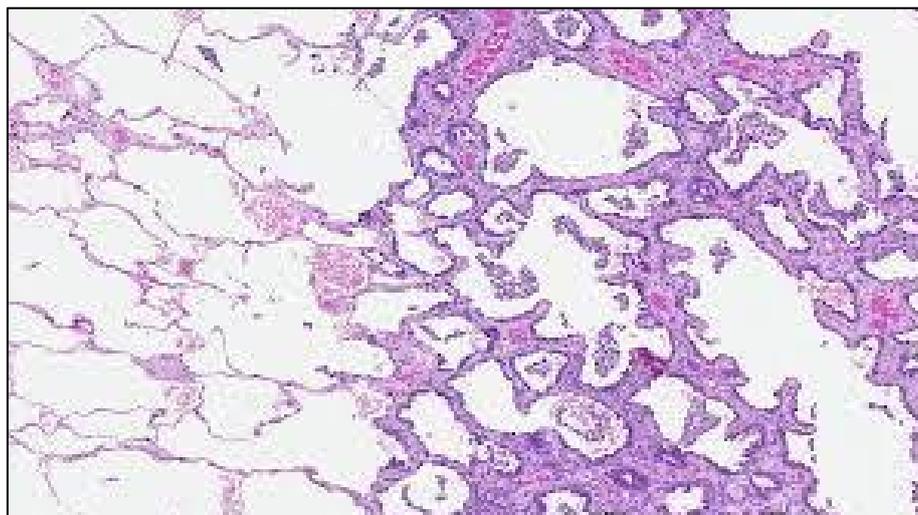


Figure 03: Vue microscopique de l'adénocarcinome [24].

1.4.1.2. Les carcinomes épidermoïdes

Le carcinome épidermoïde est une tumeur maligne épithéliale présentant des signes de différenciation malpighienne, tels que des ponts d'union intercellulaire et/ou une synthèse de kératine [25]. Il représente environ 35 % des cancers du poumon non à petites cellules (Figure 4) [23].

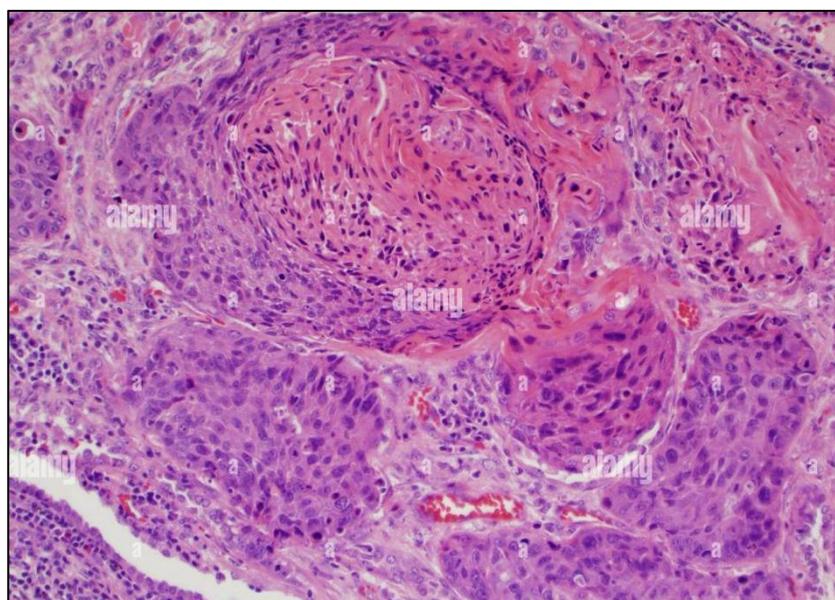


Figure 04: Vue microscopique de carcinome épidermoïde [24].

1.4.1.3. Les carcinomes à grandes cellules

Le carcinome à grandes cellules est une tumeur périphérique volumineuse souvent nécrosée [4], dépourvue de signes morphologiques de différenciation glandulaire ou squameuse. Il représente 5 à 10 % de tous les cancers du poumon non à petites cellules. Parmi ces carcinomes, on trouve le carcinome neuroendocrine et le carcinome basaloïde [26].

a. Le carcinome neuroendocrine

Est une variante des carcinomes à grandes cellules. Il est caractérisé par une architecture organoïde, insulaire, trabéculaire, ainsi que par la présence de rosettes et de palissades péri-vasculaires [27].

b. Le carcinome basaloïde

Est une forme rare de carcinome faiblement différencié à développement préférentiellement central/endobronchique. Le motif basaloïde est caractérisé par une architecture solide lobulaire, avec une disposition palissadique des cellules en périphérie (Figure 5) [24].

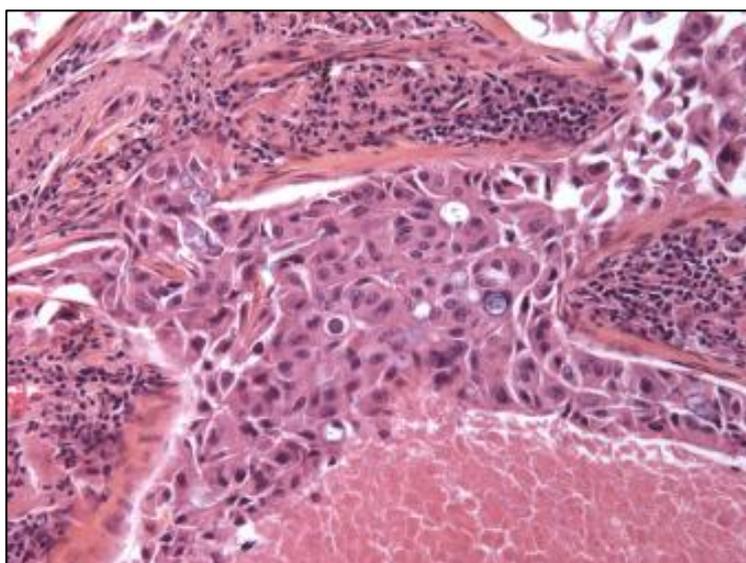


Figure 05 : Vue microscopique de carcinomes à grande cellule [24].

1.5. Les symptômes

Les symptômes les plus fréquents sont la dyspnée, les hémoptysies, la toux, les sifflements, les douleurs thoraciques ou abdominales, la dysphonie, l'hippocratisme digital, la dysphagie et les signes d'atteinte générale comme la fatigue, la perte de poids et l'inappétence [10].

1.5.1. Symptômes liés au métastases

Il est possible que des métastases révèlent également la présence d'un cancer du poumon :

- Douleur osseuse, si des métastases osseuses se produisent.
- Maux de tête, s'ils sont transférés au cerveau, convulsions, confusion, troubles de la parole, de l'équilibre et de la coordination des mouvements.
- Jaunisse, si des métastases hépatiques se produisent [9].

1.5.2. Syndromes paranéoplasiques

Les syndromes paranéoplasiques (SPN) sont définis par un ensemble de signes cliniques, radiologiques ou biologiques associés à une néoplasie, indépendamment de sa localisation. Ils peuvent être associés à tous les types de cancer, mais sont plus fréquemment rencontrés dans le cancer pulmonaire, où ils touchent 10 % des patients [28]. Parmi les nombreux syndromes paranéoplasiques endocriniens, le syndrome de Cushing est le plus fréquent, suivi de l'hypercalcémie et du Syndrome de sécrétion inappropriée d'hormone antidiurétique (SIADH) [29].

1.6. Les Facteurs de risque

Le développement de cancer du poumon est lié à plusieurs facteurs de risque qui incluent [30]:

1.6.1. Le tabagisme

a. Le tabagisme actif (cigarette, cigare, pipe, narguilé)

Le tabagisme responsable d'environ 85 à 90% des cas de cancer du poumon, dont 90 % chez les hommes et 80 % chez les femmes. Plus de 40 agents cancérigènes ont été identifiés dans la fumée de cigarette. Fumer de la marijuana et de la cocaïne augmente aussi le risque de développement de cancer du poumon (Figure 6) [31].



Figure 06: poumon d'un fumeur et d'un non fumeur [9].

b. Le tabagisme passif

L'exposition à la fumée de tabac dans l'environnement par les non-fumeurs, en particulier dans les lieux de travail, augmente le risque de développer un cancer du poumon [31].

1.6.2. Le radon

La seconde cause de cancer du poumon après le tabagisme [32]. C'est un gaz rare qui provient de la désintégration du radium. Le radon et ses produits de filiation se lient aux particules aériennes en suspension. Ils peuvent être inhalés et atteindre l'épithélium bronchique. Les particules alpha peuvent créer des lésions du double-brin de l'ADN et ainsi initier un cancer [33].

1.6.3. Les expositions professionnelles

a. L'amiante

L'amiante est le facteur de risque professionnel le plus fréquemment impliqué dans le CBP [34]. Toutes les sortes de fibres d'amiante sont aujourd'hui reconnues comme facteur de risque du cancer bronchique [35].

b. Autre produit

Il existe d'autres substances cancérigènes d'origine professionnelle que l'amiante, favorisant le cancer du poumon [35]: la silice cristalline, les hydrocarbures polycycliques aromatiques, différents métaux (arsenic, cadmium, béryllium, certains composés du chrome et du nickel) et produits chimiques (bis-chlorométhyléther) [36].

1.6.4. Pollution atmosphérique

Le cancer du poumon pourrait être l'un des effets indésirables à long terme de l'exposition cumulée à la pollution atmosphérique ambiante [37], les polluants atmosphériques pénètrent dans les poumons par le système respiratoire et peuvent augmenter le risque de cancer du poumon [38].

1.6.5. Les antécédents personnels et familiaux

Les antécédents personnels d'affections respiratoires sont également un facteur de risque de cancer du poumon, en particulier les personnes atteintes de tuberculose. De même, les personnes qui ont déjà été atteintes d'un cancer pulmonaire ont davantage de risque de développer un deuxième cancer pulmonaire que la population générale. Par ailleurs, un cas de cancer pulmonaire familial augmente également le risque de survenue de ce cancer [35].

1.7. Les stades de cancer du poumon

La stadification initiale est un élément essentiel lors de la prise en charge d'un cancer broncho-pulmonaire. Elle repose sur la classification TNM [39]. Le bilan d'extension du CBP précise le stade de la tumeur : I à IV en fonction de la taille de la tumeur, de l'atteinte ganglionnaire médiastinale, de l'envahissement d'organes par contiguïté et de l'atteinte métastatique [40].

1.7.1. La classification TNM

La 8^{ème} édition de la classification TNM (Tableaux 1 et 2), proposée par l'IASLC devrait entrer en vigueur en janvier 2017, après avoir été validée par l'Union internationale contre le cancer (UICC) [40].

Tableau 01 : Répartition par stade selon la 8^{ème} classification TNM pour les cancers pulmonaires [41] .

Stade	T	N	M
Cancer occulte	Tx	0	0
0	Tis	0	0
IA1	T1a(mi)	0	0
	T1a	0	0
IA2	TAb	0	0
IA3	T1c	0	0
IB	T2a	0	0
IIB	T1a-c	1	0
	T2a	1	0
	T2b	1	0
	T3	0	0
IIIA	T1a-c	2	0
	T2a-b	2	0
	T3	1	0
	T4	0	0
IIIB	T4	1	0
	T1a-c	3	0
	T2a-b	3	0
	T3	2	0
IIIC	T4	2	0
	T3	3	0
IVA	T4	3	0
	Tout T	Tout N	M1a
IVB	Tout T	Tout N	M2b
	Tout T	Tout N	M1c

Tableau 02: Proposition des descripteurs T,N et M pour la huitième édition de la classification TNM pour le cancer du poumon [42].

T: tumeur primaire	
Tx	la tumeur primaire ne peut pas être évaluée ou tumeur prouvée par la présence de cellules malignes dans les expectorations ou les lavages bronchiques mais non visualisée par imagerie ou bronchoscopie
T0	Aucun signe de tumeur primitive
Tis	carcinome in situ
T1	Tumeur ≤ 3 cm dans sa plus grande dimension entourée de poumon ou de plèvre viscérale sans preuve bronchoscopique d'invasion plus proximale que la bronche lobaire (c'est-à-dire pas dans la bronche principale)
T1a(mi)	Adénocarcinome mini-invasif
T1a	La tumeur est ≤ 1 cm dans sa plus grande dimension
T1b	Tumeur > 1 cm mais ≤ 2 cm dans la plus grande dimension
T1c	Tumeur > 2 cm mais ≤ 3 cm dans la plus grande dimension
Tumeur T2	> 3 cm mais ≤ 5 cm ou tumeur présentant l'une des caractéristiques suivantes : - Implique la bronche principale quelle que soit la distance de la carène mais sans implication de la carène - Envahit la plèvre viscérale - Associée à une atelectasie ou une pneumonie obstructive qui s'étend à la région hilare, impliquant une partie ou la totalité du poumon
T2a	Tumeur > 3 cm mais ≤ 4 cm dans la plus grande dimension
T2b	Tumeur > 4 cm mais ≤ 5 cm dans la plus grande dimension
Tumeur T3	> 5 cm mais ≤ 7 cm dans sa plus grande dimension ou associée à un ou plusieurs nodules tumoraux séparés dans le même lobe que la tumeur primaire ou envahissant directement l'une des structures suivantes : paroi thoracique (y compris la plèvre pariétale et les tumeurs du sillon supérieur) , nerf phrénique, péricarde pariétal
Tumeur T4	> 7 cm dans sa plus grande dimension ou associée à un ou plusieurs nodules tumoraux distincts dans un lobe ipsilatéral différent de celui de la tumeur primitive ou envahissant l'une des structures suivantes : diaphragme, médiastin, cœur, gros vaisseaux, trachée, nerf laryngé récurrent , œsophage, corps vertébral et carène
N : atteinte des ganglions lymphatiques régionaux	
N0	Aucune métastase ganglionnaire régionale
Nx	Les ganglions lymphatiques régionaux ne peuvent pas être évalués
N1	Métastases dans les ganglions lymphatiques péribronchiques et/ou hilaires homolatéraux et les ganglions intrapulmonaires, y compris atteinte par extension directe

N2	Métastases dans les ganglions lymphatiques médiastinaux et/ou sous-carinaux homolatéraux
N3	Métastases dans le médiastin controlatéral, hilair controlatéral, scalène ipsilatéral ou controlatéral ou ganglion(s) lymphatique(s) supraclaviculaire(s)
M : métastase à distance	
M0	Pas de métastase à distance
M1	Présence de métastases à distance
M1a	Nodule(s) tumoral(s) séparé(s) dans un lobe controlatéral ; tumeur avec nodule(s) pleural ou péricardique ou épanchement pleural ou péricardique malin.
M1b	Métastase extrathoracique unique
M1c	Multiplés métastases extrathoraciques dans un ou plusieurs organes

1.8. Diagnostic et traitement de cancer du poumon

1.8.1. Diagnostic

1.8.1.1. Examen clinique

L'examen clinique a pour objectifs de faire un bilan de l'état général à l'aide de différents moyens, d'identifier les antécédents médicaux et les pathologies dont le patient souffre, de recenser les traitements en cours, les facteurs de risque et sa dépendance au tabac, s'il fume [9].

1.8.1.2. Examen complémentaire

a) Biologie (examens systématiques)

Les examens biologiques systématiques sont: l'hémogramme, l'ionogramme complet, la clairance de la créatinine avant le scanner injecté, la calcémie, le bilan hépatique (transaminases, gamma GT, bilirubine), le TP et la TCA [4].

b) Imagerie

Une imagerie de première intention doit être pratiquée rapidement devant tous symptômes ou contexte évocateurs de cancer bronchique [4].

➤ Radiographie du thorax (face et profil)

C'est un examen facile qui permet une orientation rapide, ses performances, sa sensibilité et sa spécificité sont limitées [4]. L'objectif de cet examen est de révéler la présence d'anomalies dans les poumons [9].

➤ Tomodensitométrie thoracique (TDM)

La TDM joue un rôle fondamental dans l'exploration des cancers broncho-pulmonaires, car elle permet d'affirmer sa localisation thoracique, d'évoquer la suspicion de malignité, et offre des possibilités de caractérisation des nodules ou des masses pulmonaires [4].

➤ Le scanner TEP-TDM

La tomographie par émission de positons TEP associée à une TDM est une modalité d'imagerie de plus en plus répandue. Elle permet une exploration non invasive, morphologique et métabolique du patient, grâce à un traceur [43]. Actuellement le traceur le plus utilisé est le fluorodésoxyglucose (18FDG), c'est un analogue du glucose, porteur d'un atome radioactif, qui va se fixer sur les tissus qui présentent un hypermétabolisme glucidique (hyperfixation), comme les tissus cancéreux [44].

➤ Imagerie par résonance magnétique (IRM)

Imagerie par résonance magnétique (IRM) est un examen d'imagerie médicale qui aide au diagnostic du cancer. Une IRM permet d'obtenir des images précises de

l'intérieur du corps grâce à l'utilisation d'un aimant et d'ondes électromagnétiques [45]. Elle est la technique de choix pour évaluer l'extension des tumeurs de l'apex pulmonaire [46].

c) Endoscopie bronchique (Fibroskopie bronchique)

L'endoscopie bronchique est un examen incontournable pour le diagnostic de cancer bronchique [47]. Il s'agit essentiellement de l'extraction des corps étrangers bronchiques. Elle peut être réalisée à l'aide d'une bronchoscopie rigide ou souple [48].

1.8.2. Traitement

Les options de traitement de cancer du poumon comprennent généralement la résection chirurgicale, la radiothérapie, la chimiothérapie et la thérapie ciblée [49].

1.8.2.1. La chirurgie

La chirurgie d'exérèse est un traitement du CBNPC. Elle est le traitement de choix dans les cancers de stade I et II et garde une place importante pour les cancers de stade IIIA dans le cadre d'un traitement multi-modal [50]. Il existe trois approches chirurgicales validées : la pneumonectomie (ablation d'un poumon entier), la lobectomie (ablation de tout le lobe contenant la tumeur), et la segmentectomie (ablation partielle d'un lobe) [51].

1.8.2.2. La chimiothérapie

Le traitement chimiothérapeutique peut être soit localisé après une intervention chirurgicale, soit systémique. Le choix du traitement est adapté en fonction des résultats anatomiques et cliniques postopératoires. La chimiothérapie implique généralement l'utilisation d'une combinaison de sels de platine tels que le carboplatine, le cisplatine ou l'oxaliplatine, ainsi que d'inhibiteurs des voies de signalisation affectées par les mutations, telles que EGFR, ALK, KRAS ou BRAF1 [8].

1.8.2.3. la radiothérapie

La radiothérapie, seule ou combinée à une chimiothérapie, joue un rôle majeur dans le traitement des CBNPC localement avancés [52]. La radiothérapie en conditions stéréotaxiques est une technique d'irradiation validée pour les tumeurs primitives du poumon localisées de stade précoce : c'est un traitement de référence pour les patients inopérables et une option réelle chez les patients opérables ou ceux qui refusent la chirurgie [51].

1.8.2.4. La thérapie ciblée

Les thérapies moléculaires ciblées représentent une nouvelle classe thérapeutique dans la stratégie de traitement des cancers. Le principe général guidant leur développement est d'identifier une altération oncogénique dans les cellules tumorales

puis de la cibler spécifiquement, actuellement essentiellement avec un anticorps monoclonal ou un inhibiteur pharmacologique [53]. Les médicaments les plus souvent employés comme traitement ciblé du CBNPC sont : Erlotinib (Tarceva) anti EGFR , Bevacizumab (Avastin) anti VEGF, Gefitinib (Iressa) anti EGFR [4].

1.9. Prévention

Environ 85% de tous les décès dus au cancer du poumon sont estimés être attribuables au tabagisme. Éviter l'utilisation du tabac est la mesure la plus efficace pour prévenir le cancer du poumon. L'agence pour la recherche et la qualité des soins de santé a élaboré un ensemble de directives de sevrage tabagique clinique. Les six éléments majeurs des lignes directrices comprennent ce qui suit :

- Réduire l'accès des mineurs aux produits du tabac.
- Diffuser des programmes de prévention scolaires efficaces et des stratégies médiatiques.
- Augmenter le coût des produits du tabac en augmentant les taxes.
- Utiliser les taxes sur les produits du tabac pour financer des interventions au niveau de la communauté, y compris les médias de masse.
- Fournir des stratégies éprouvées pour arrêter de fumer par l'intermédiaire des organisations de soins de santé.
- Adopter des lois et des politiques sans fumée... [54], et d'autres mesures de prévention comprennent des choix alimentaires sains et le maintien d'un mode de vie physiquement actif [55].

Matériels et Méthodes

2. Méthodes

2.1. L'objectif de travail

L'objectif de notre travail est d'analyser l'expression des gènes différentiels liés au cancer du poumon à l'aide d'outils bioinformatiques.

2.2. Méthode d'analyse

2.2.1. Source des données et informations sur les puces à ADN

Dans cette étude, nous avons utilisé des données des puces à ADN téléchargées à partir du référentiel public de données de génomique fonctionnelle appelée la base de données *Gene Expression Omnibus* (GEO) (<https://www.ncbi.nlm.nih.gov/gds>). Nous avons sélectionné trois profils d'expression génique (GSE17558, GSE21933, GSE68465) pour étudier le cancer du poumon. Les données de la matrice de GSE17558 incluent 8 échantillons de tissus de CBP et 8 échantillons de tissus normaux, basés sur la plateforme GPL6883 (*Illumina HumanRef-8 v3.0 expression beadchip*). Les données de la matrice GSE21933 comprennent 21 échantillons de tissus de CBP et 21 échantillons de tissus normaux, basés sur la plateforme GPL6254 (*Phalanx Human OneArray*). Concernant les données de la matrice GSE68465, ils comprennent 223 échantillons de tissus de CBP et 19 échantillons de tissus normaux, basés sur la plateforme GPL96 (*[HG-U133A] Affymetrix Human Genome U133A Array*).

2.2.2. Traitement des données et identification des DEG

Les trois profils d'expression d'ARNm identifiés dans la base de données GEO ont été soumis à l'analyse en utilisant les outils en ligne de GEO2R (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>). Un $|\log_2\text{-fold change (FC)}| > 1.5$ et une valeur $P < 0,05$ ont été utilisés comme critères de sélection pour l'identification des gènes différentiellement exprimés (DEG), les données au format TXT ont été traitées à l'aide de l'outil Venn Diagram (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) pour identifier les DEG qui sont communs aux trois profils. Les DEG avec un $\log FC > 0$ ont été identifiés comme des gènes surexprimés (*upregulated genes*), alors que les DEG avec un $\log FC < 0$ ont été considérés comme des gènes sous-exprimés (*down-regulated genes*).

2.2.3. Étude de l'ontologie des gènes et de l'enrichissement des voies

L'approche fréquemment utilisée dans les études fonctionnelles de données transcriptomiques ou génomiques à haut débit est l'analyse de l'ontologie des gènes (*Gene Ontology*, GO). Cette analyse implique l'utilisation d'un outil bioinformatique en ligne appelé le logiciel DAVID (*version 6.8*, <https://david.ncifcrf.gov/>) pour effectuer une analyse d'enrichissement des voies GO (*Biological process* : BP, *Cellular component* : CC, et *Molecular function* : MF) ainsi que du *Kyoto Encyclopedia of Genes and Genomes* (KEGG). Le critère de sélection retenu était un seuil de $P < 0,05$.

2.2.4. Utilisation de la PPI pour l'analyse des réseaux et des modules

Afin d'étudier les interactions entre les DEG, nous avons utilisé la base de données en ligne STRING (version 11.5, <https://string-db.org/>). Un score d'interaction minimum de confiance élevé (0,900) a été choisi pour l'analyse. Ensuite, après avoir éliminé les nœuds déconnectés, nous avons visualisé un réseau de Protein-Protein Interaction (PPI) à l'aide du logiciel Cytoscape 3.9.0. De plus, le plug-in MCODE de Cytoscape a été utilisé pour identifier les modules du réseau PPI, en appliquant un seuil de degré de 2, un seuil de score de nœud de 0,2, un k-core de 2 et une profondeur maximale de 100.

2.2.5. Sélection des gènes clés

Nous avons utilisé le plug-in cytoHubba du logiciel Cytoscape pour extraire les gènes pertinents. Ensuite, nous avons sélectionné les huit principaux gènes clés classés selon les algorithmes topologiques *Maximal Clique Centrality* (MCC) et *Degree*.

2.2.6. Analyse de survie globale et validation de l'expression des gènes clés

Nous avons utilisé l'outil *Kaplan Meier Plotter* (<http://kmplot.com/analysis/index.php?p=service>) pour évaluer l'impact de huit gènes clés sur la survie globale (OS) des patients atteints de CBP. Pour estimer le taux de survie globale des patients atteints de cancer bronchique, les échantillons ont été divisés en groupes à expression élevée et faible en fonction de la médiane d'expression. Le rapport de risque (*Hazard Ratio*, HR), les intervalles de confiance à 95 % correspondants et les valeurs P du log-rank ont été calculés et présentés sur le graphique.

Ensuite, l'outil *Gene Expression Profiling Interactive Analysis* GEPIA2 (<http://gepia2.cancer-pku.cn/>) a été utilisé pour analyser les données d'expression de séquençage d'ARN des patients atteints de CBP en les comparant à des échantillons témoins (avec une valeur de coupure de $|\text{Log}_2\text{FC}| = 1$ et un seuil de p-valeur de 0,05).

Résultats et discussion

3. Résultats et discussion

3.1. Résultats

3.1.1. Analyse des données

Les trois profils d'expression d'ARNm provenant de la base de données GEO ont été soumis à une analyse à l'aide de l'outil en ligne GEO2R en utilisant une valeur $P < 0,05$ et $|\log_2\text{-fold change (FC)}| > 1.5$. L'ensemble de données GSE17558 contient 2133 DEG, dont 979 gènes surexprimés et 1154 gènes sous-exprimés. L'ensemble de données GSE21933 contient 796 DEG, comprenant 353 gènes surexprimés et 443 gènes sous-exprimés. Enfin, l'ensemble de données GSE68465 présente 987 DEG, avec 524 gènes surexprimés et 463 gènes sous-exprimés. Les tracés volcaniques (*Volcano Plot*) des trois ensembles de données GEO sont présentés dans la figure 7.

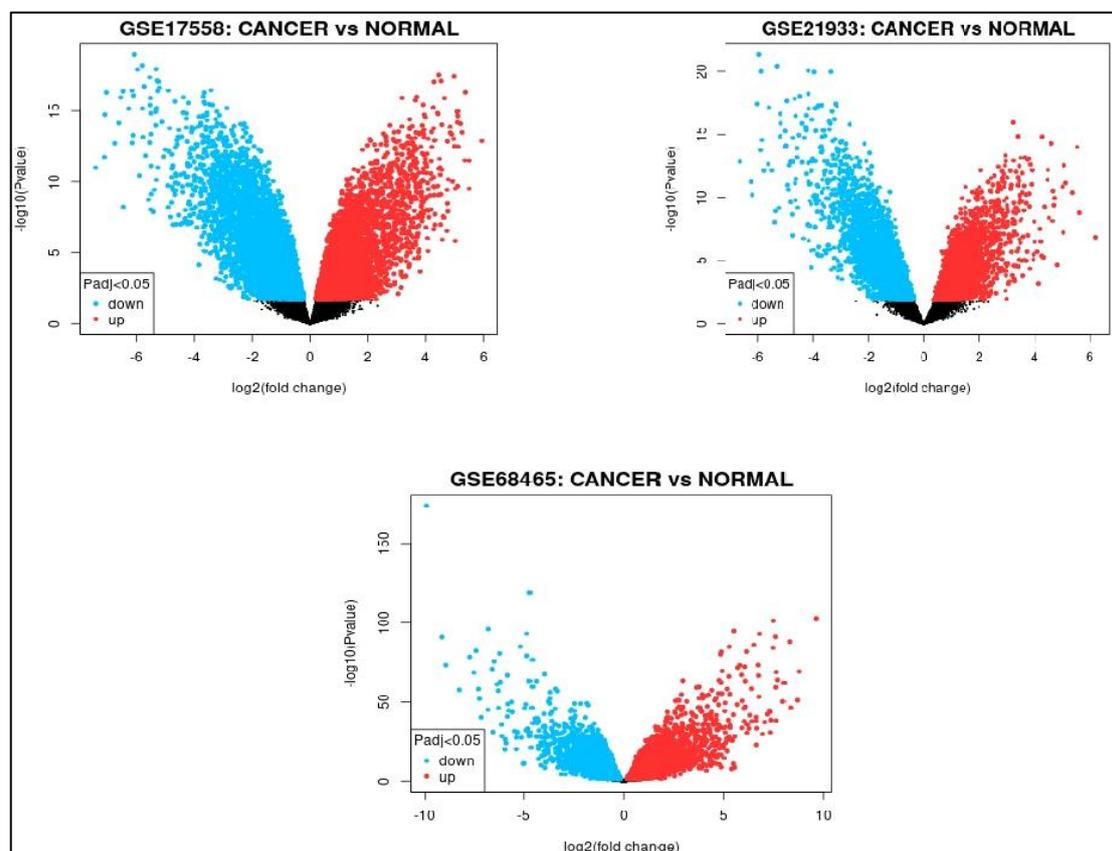


Figure 07: Les tracés volcaniques (*Volcano plots*) des trois ensembles de données GEO, le bleu représente les gènes sous-exprimés, le rouge représente les gènes surexprimés ($P < 0:05$ et $\log |FC| > 1.5$).

À l'aide de la base de données Venn Diagram, nous avons identifié les DEG communs aux trois profils. La figure 8 montre que 179 DEG ont été trouvés dans les trois ensembles de données GEO, comprenant 78 gènes surexprimés et 101 gènes sous-exprimés. La liste des DEG communs surexprimés et sous-exprimés est montrée dans le tableau 3.

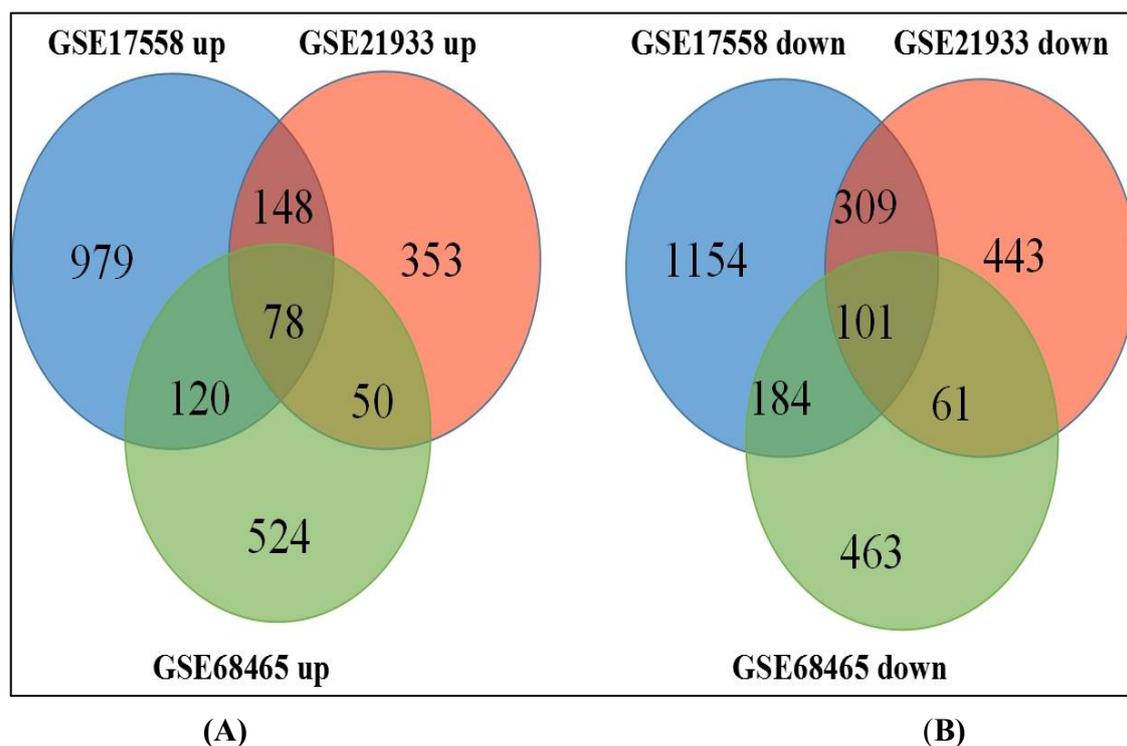


Figure 08: Les DEG communs aux trois *datasets*. (A) les 78 gènes surexprimés. (B) les 101 gènes sous-exprimés.

Tableau 03: Liste des DEG communs aux trois *datasets*.

DEG communes	Gènes
Surexprimés	TPX2 CCNB1 BIRC5 FOXM1 CDC6 AURKA KIF14 AURKB PITX1 KIF4A CA9 BMP7 MCM10 TYMS MELK ZWINT NDC80 OIP5 CCNA2 GAL COCH KIF23 CENPE CCNB2 PRC1 CDT1 CEP55 MPP6 MCM2 MCM4 TFAP2A MKI67 EXO1 CENPI KIAA0101 SLC16A1 NEK2 GINS2 ESPL1 C16orf59 PLK1 FBN2 EZH2 CHEK1 KIF11 SKP2 CDH2 ZWILCH SMC4 LEF1 NCAPG2 KIF2C CENPN PLK4 BUB1 PBK TRIP13 ASPM POLE2 PSAT1 RRM2 FEN1 HELLS FANCI RAD54L WDHD1 KIF18A KIF15 DKK1 RNASEH2A HMMR ERCC6L KIF20A CDKN3 CENPF NUSAP1 KRT23
Sous-exprimés	CRTAM LILRB2 FGR AQP9 FXVD6 FRY CSF3R PAPSS2 MAOA LAMP3 SERPING1 FCN1 ACADL PILRA RHOTB2 RRAD SFTPD OLR1 ADH1C RUNX1T1 CD93 MSR1 TPPP3 SCNN1B CLIC5 FHL5 NR4A3 NKX2-1 TNFSF13 RPS6KA2 SCTR NR3C2 SCGB1A1 ACP5 ALOX5AP ABCA6 RASL12 TLR7 TMC5 CTSE TRPC6 KRT7 FLRT3 IL10RA SFTPC MYL9 RORC FOSB MARCO C4BPA PRELP TREM1 IL18R1 HSD17B6 CYP4B1 HBEGF MUC1 CD36 AQP4 COL4A3 SELE NKG7 STEAP4 AOC3 GJA4 MYH11 CPM IRAK3 VSIG4 SRGN HPGD SCEL TMPRSS2 DUOX1 DPT MMP28 TYROBP ALOX15B PTGDS DMBT1 LIMCH1 CD52 SFTPB COX7A1 IL7R CCL21 FMO2 LMO3 IL33 HLA-DRB1 FCN3 CRTAC1 FAM107A CTSH KCNJ15 C1orf116 CRISPLD2 RND1 CLDN18 CDH5

3.1.2. Étude de l'ontologie génique et de l'enrichissement des voies

L'analyse d'enrichissement des voies de GO a été effectuée sur les DEG communs (les gènes surexprimés et sous-exprimés). Le critère de sélection pour l'analyse été une valeur $P < 0,05$. L'analyse a compris *Biological process* (BP), *Cellular component* (CC), *Molecular function* (MF).

Dans les BP, on observe que les gènes surexprimés ont été principalement enrichis dans les voies suivantes : *cell division*, *mitotic cell cycle*, *mitotic sister chromatid segregation*, *mitotic spindle assembly checkpoint*, *mitotic spindle organization*, *G2/M transition of mitotic cell cycle*. En revanche, les gènes sous-exprimés ont été principalement enrichis dans des voies suivantes : *negative regulation of T cell proliferation*, *receptor-mediated endocytosis*, *response to lipopolysaccharide*, *female pregnancy*, *negative regulation of fatty acid oxidation*, *cytokine-mediated signaling pathway*.

Dans les CC, les résultats de l'analyse ont montré que les gènes surexprimés ont été principalement enrichis dans les voies suivantes : *kinetochore*, *spindle*, *midbody*, *nucleoplasm*, *nucleus*, *chromosome*, *centromeric region*. Pour les gènes sous-exprimés

ont été principalement enrichis dans les voies suivantes : *plasma membrane, extracellular space, integral component of membrane, extracellular region, integral component of plasma membrane, collagen trimer*.

Dans les MF, les gènes surexprimés ont été principalement enrichis dans les voies: *microtubule binding, ATP binding, microtubule motor activity, protein binding, ATPase activity, protein kinase binding*. Les gènes sous-exprimés ont été enrichis dans les voies suivantes: *scavenger receptor activity, pattern recognition receptor activity, carbohydrate binding, flavin adenine dinucleotide binding, receptor binding, beta-amyloid binding*. L'analyse GO des DEG communs est présentée dans le tableau 4 .

Tableau 04: Analyse GO des DEG communs dans le cancer du poumon.

Catégorie	Terme	N	%	Valeur de P
Les gènes surexprimés				
GOTERM_BP_DIRECT	<i>cell division</i>	26	34,2	2,7E-24
GOTERM_BP_DIRECT	<i>mitotic cell cycle</i>	16	21,1	1,3E-17
GOTERM_BP_DIRECT	<i>mitotic sister chromatid segregation</i>	10	13,2	1,0E-14
GOTERM_BP_DIRECT	<i>mitotic spindle assembly checkpoint</i>	9	11,8	3,4E-13
GOTERM_BP_DIRECT	<i>mitotic spindle organization</i>	9	11,8	4,9E-11
GOTERM_BP_DIRECT	<i>G2/M transition of mitotic cell cycle</i>	9	11,8	4,9E-11
GOTERM_CC_DIRECT	<i>kinetochore</i>	16	21,1	1,1E-17
GOTERM_CC_DIRECT	<i>spindle</i>	14	18,4	6,4E-15
GOTERM_CC_DIRECT	<i>midbody</i>	14	18,4	1,3E-13
GOTERM_CC_DIRECT	<i>nucleoplasm</i>	42	55,3	8,8E-12
GOTERM_CC_DIRECT	<i>nucleus</i>	51	67,1	1,7E-11
GOTERM_CC_DIRECT	<i>chromosome, centromeric region</i>	9	11,8	7,9E-11
GOTERM_MF_DIRECT	<i>microtubule binding</i>	15	19,7	1,5E-12
GOTERM_MF_DIRECT	<i>ATP binding</i>	27	35,5	5,5E-11
GOTERM_MF_DIRECT	<i>microtubule motor activity</i>	9	11,8	8,1E-11
GOTERM_MF_DIRECT	<i>protein binding</i>	72	94,7	8,4E-10
GOTERM_MF_DIRECT	<i>ATPase activity</i>	12	15,8	6,5E-7
GOTERM_MF_DIRECT	<i>protein kinase binding</i>	11	14,5	3,1E-5
les gènes sous-exprimées				
GOTERM_BP_DIRECT	<i>negative regulation of T cell proliferation</i>	5	4,9	1,2E-4
GOTERM_BP_DIRECT	<i>receptor-mediated endocytosis</i>	5	4,9	7,1E-4

GOTERM_BP_DIRECT	<i>response to lipopolysaccharide</i>	6	5,8	8,5E-4
GOTERM_BP_DIRECT	<i>female pregnancy</i>	5	4,9	9,0E-4
GOTERM_BP_DIRECT	<i>negative regulation of fatty acid oxidation</i>	3	2,9	9,3E-4
GOTERM_BP_DIRECT	<i>cytokine-mediated signaling pathway</i>	6	5,8	1,3E-3
GOTERM_CC_DIRECT	<i>plasma membrane</i>	51	49,5	4,5E-7
GOTERM_CC_DIRECT	<i>extracellular space</i>	28	27,2	5,8E-7
GOTERM_CC_DIRECT	<i>integral component of membrane</i>	51	49,5	6,5E-7
GOTERM_CC_DIRECT	<i>extracellular region</i>	28	27,2	3,0E-6
GOTERM_CC_DIRECT	<i>integral component of plasma membrane</i>	22	21,4	6,3E-6
GOTERM_CC_DIRECT	<i>collagen trimer</i>	7	6,8	6,4E-6
GOTERM_MF_DIRECT	<i>scavenger receptor activity</i>	6	5,8	7,2E-6
GOTERM_MF_DIRECT	<i>pattern recognition receptor activity</i>	4	3,9	1,1E-4
GOTERM_MF_DIRECT	<i>carbohydrate binding</i>	6	5,8	4,4E-3
GOTERM_MF_DIRECT	<i>flavin adenine dinucleotide binding</i>	4	3,9	5,6E-3
GOTERM_MF_DIRECT	<i>receptor binding</i>	8	7,8	5,9E-3
GOTERM_MF_DIRECT	<i>beta-amyloid binding</i>	4	3,9	1,0E-2

3.1.3. Étude de Kyoto Encyclopedia of Genes and Genomes (KEGG)

Le critère de sélection de l'analyse KEGG était la valeur $P < 0,05$. Les résultats ont identifié huit voies de signalisation. Les voies de signalisation associées aux DEG étaient principalement enrichies dans les voies suivantes : *Cell cycle*, *DNA replication*, *Progesterone-mediated oocyte maturation*, *Oocyte meiosis* pour les gènes surexprimés et *Phagosome*, *Drug metabolism-cytochrome P450*, *Cytokine-cytokine receptor interaction*, *Tyrosine metabolism* pour les gènes sous-exprimés. L'analyse de la voie KEGG des DEG commun est présentée dans le tableau 5.

Tableau 05 : Analyse KEGG des DEG courants dans le cancer des poumons.

Catégorie	Terme	Nm	%	Valeur de P
Les gènes surexprimés				
KEGG_PATHWAY	<i>Cell cycle</i>	11	14,5	3,8E-11
KEGG_PATHWAY	<i>DNA replication</i>	5	6,6	1,2E-5
KEGG_PATHWAY	<i>Progesterone-mediated oocyte maturation</i>	6	7,9	4,8E-5
KEGG_PATHWAY	<i>Oocyte meiosis</i>	6	7,9	1,6E-4
les gènes sous-exprimées				
KEGG_PATHWAY	<i>Phagosome</i>	6	5,8	6,1E-3
KEGG_PATHWAY	<i>Drug metabolism - cytochrome P450</i>	4	3,9	1,8E-2
KEGG_PATHWAY	<i>Cytokine-cytokine receptor interaction</i>	7	6,8	2,5E-2
KEGG_PATHWAY	<i>Tyrosine metabolism</i>	3	2,9	3,1E-2

3.1.4. Analyse de l'interaction protéine-protéine

Afin d'évaluer plus précisément les interactions entre les DEG communs identifiés, nous avons créé un réseau PPI en utilisant le site web STRING. Ce réseau PPI de DEG communs est présenté dans la figure 9. Nous avons sélectionné 177 nœuds (*nods*) et 1963 bords (*edges*) pour construire un réseau PPI comprenant 78 gènes surexprimés et 101 gènes sous-exprimés. Ensuite, nous avons utilisé le plug-in MCODE de Cytoscape (seuil de degré = 2, seuil de score de nœud = 0,2, k-core = 2 et profondeur max = 100) pour identifier un ensemble de huit gènes (ASPM, BUB1, AURKb, CCNA2, KIF11, CCNB1, CCNB2, NDC80).

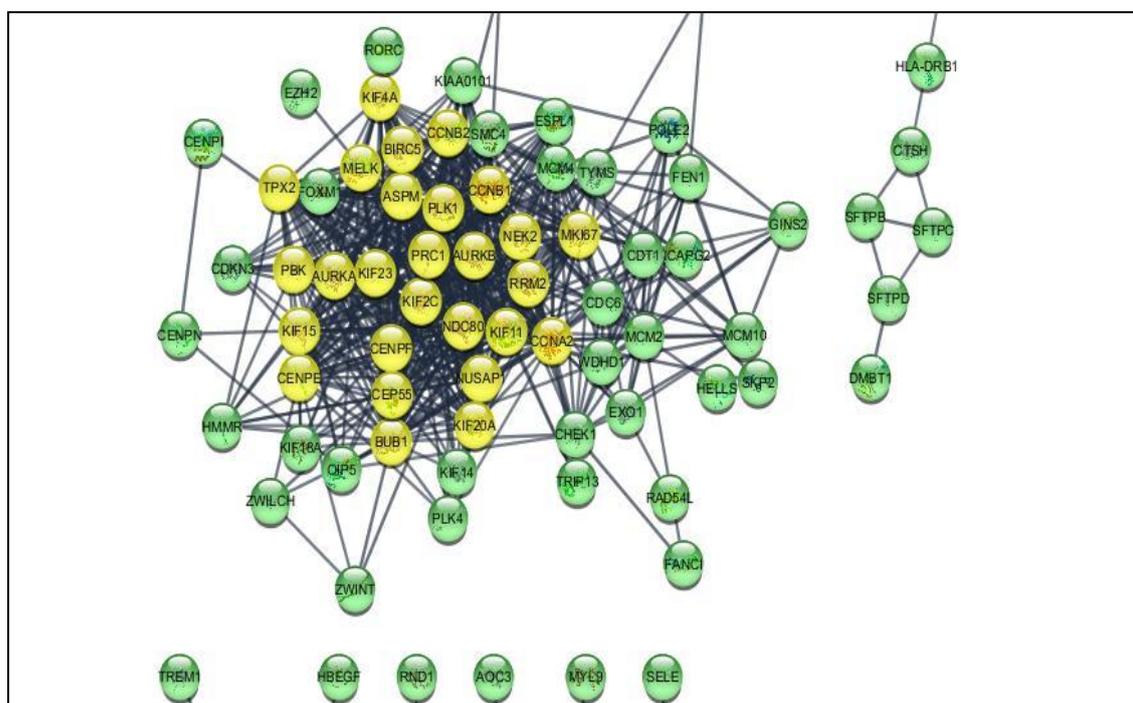


Figure 09 : Le réseau PPI, les gènes surexprimés en jaune, les gènes sous-exprimés en vert.

3.1.5. Sélection des gènes clés

Grâce à Cytohubba, une unité fonctionnelle du réseau PPI contenant huit gènes clés a été identifiée : ASPM, BUB1, AURKb, CCNA2, KIF11, CCNB1, CCNB2 et NDC80 (Figure 10). Ces gènes sont considérés comme des gènes clés associés au cancer bronchique.

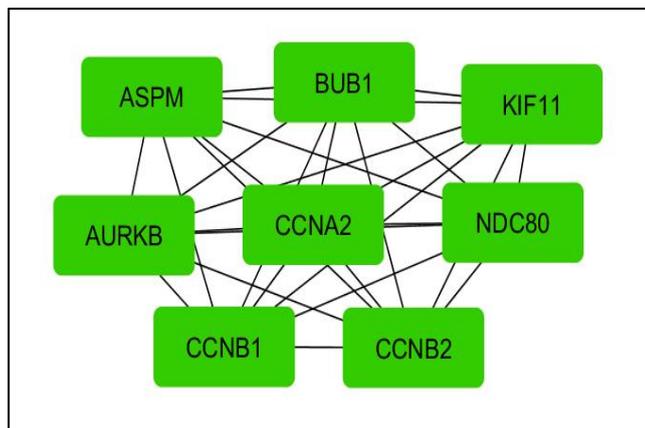


Figure 10: Les huit gènes clés classés par les algorithmes topologiques *Maximal Clique Centrality* et *degree* de Cytoscape.

3.1.6. Analyse de survie globale et validation de l'expression des gènes clés

Les résultats d'analyse de la survie globale des gènes clés ont indiqué que les niveaux d'expression de tous les huit gènes: ASPM, BUB1, AURKB, CCNA2, KIF11, CCNB1, CCNB2, NDC80 sont associés à une mauvaise survie chez les patients atteints de cancer bronchique (Figure 11).

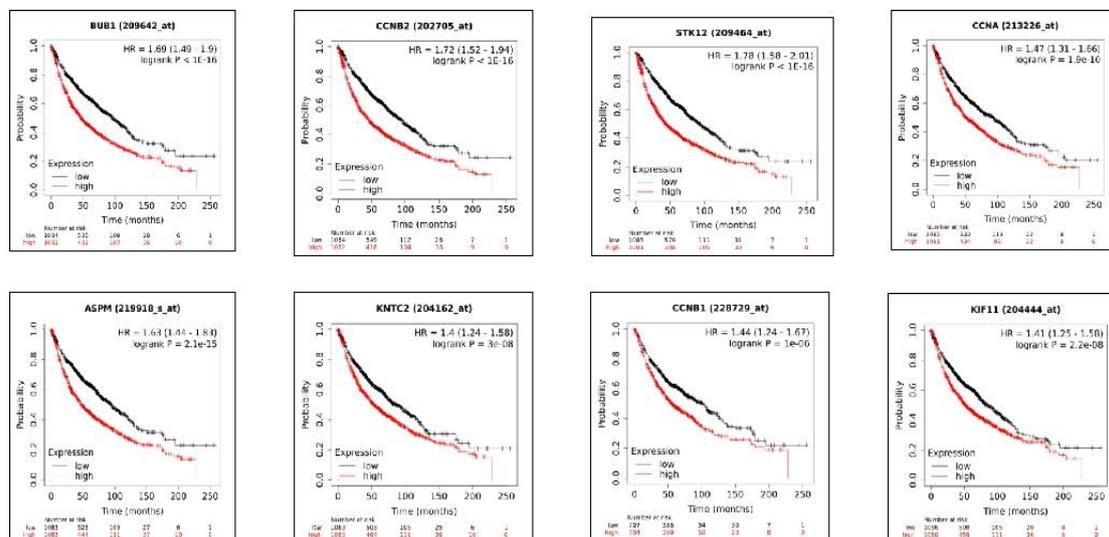


Figure 11: Analyse de la survie des huit gènes exprimés dans le cancer des poumons. Les tissus cancéreux en rouge, les tissus normaux en gris.

Les résultats obtenus à partir de la base de données GEPIA2 ont révélé une différence d'expression significative pour les gènes clés: ASPM, BUB1, AURKB, CCNA2, KIF11, CCNB1, CCNB2 et NDC80 entre les tissus tumoraux et les tissus normaux (Figure 12).

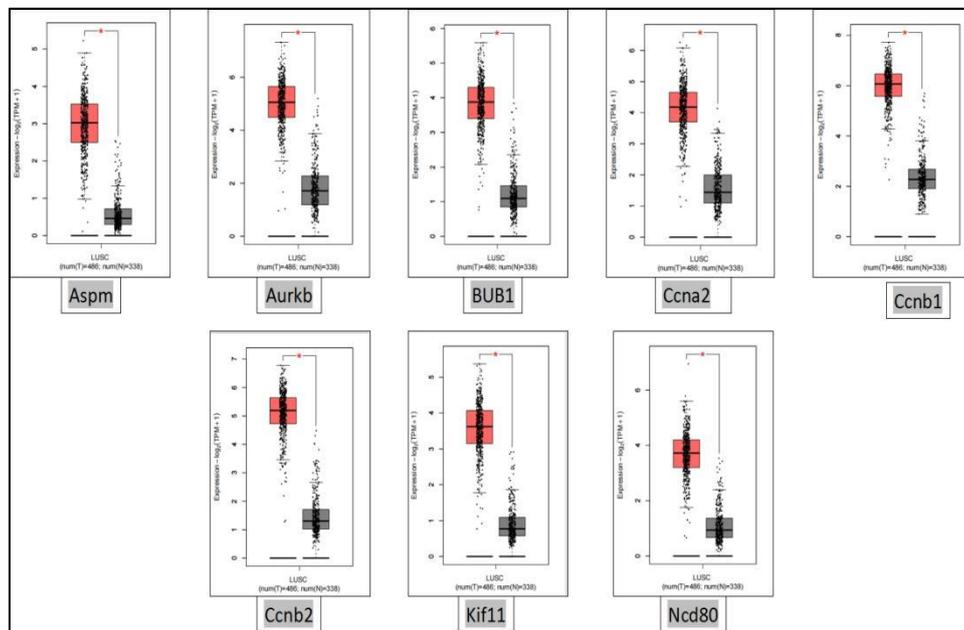


Figure 12: La vérification de l'expression des gènes clés dans les tissus de cancer des poumons par l'outil GEPIA2,(LUSC: *lung squamous cell carcinoma*).

3.2. Discussion

Le cancer des poumons est une tumeur maligne caractérisée par une croissance cellulaire incontrôlée dans les tissus du poumon [56]. C'est le deuxième type de cancer le plus courant et la principale cause de décès par le cancer dans le monde. L'adénocarcinome pulmonaire (AD) représente environ 40 % des cas de cancer du poumon, ce qui en fait le sous-type histologique le plus courant de cancer du poumon [57].

Les options de traitement pour le CBP comprennent la chirurgie, la radiothérapie, la chimiothérapie et la thérapie ciblée [58]. La thérapie ciblée est de plus en plus utilisée pour traiter les patients atteints des adénocarcinomes pulmonaires [57]. Cependant, plusieurs gènes et voies sont connus pour être impliqués, entraînant des comportements biologiques complexes. Par conséquent, l'amélioration de notre compréhension des mécanismes moléculaires du CBNPC, en particulier grâce à l'utilisation d'analyses génomiques, transcriptomiques, protéomiques et métabolomiques, devrait conduire au développement de meilleures stratégies diagnostiques et thérapeutiques basées sur l'identification de nouveaux biomarqueurs [59].

Dans cette étude, nous avons trouvé que les gènes surexprimés dans le cancer des poumons étaient principalement enrichis dans plusieurs voies telles que : *G2/M transition of mitotic cell cycle, chromosome, centromeric region, microtubule binding*. Madelaine et al. ont montré que la transition G2/M et la mitose (*G2/M transition of mitotic cell cycle*) sont des étapes du cycle cellulaire dont les gènes de régulation peuvent être altérés dans le cancer bronchique [60]. Masuda et al. ont montré que l'anomalie chromosomique est l'une des caractéristiques distinctives des cellules néoplasiques, et la présence persistante d'instabilité chromosomique a été démontrée dans les cancers humains, y compris le cancer du poumon [61]. En outre; il a été démontré que l'instabilité chromosomique prédominante dans les cellules tumorales découle de plusieurs voies différentes et est probablement régulée par des centaines de gènes [62]. Les agents de liaison aux microtubules (*microtubule-binding agents*) sont l'une des familles d'agents anticancéreux les plus anciennes et les plus diversifiées. Ces agents agissent principalement en réduisant la dynamique des microtubules, en induisant l'arrêt mitotique et la mort des cellules tumorales [63]. L'expression de la β -tubuline de classe III émerge comme un marqueur de résistance aux taxanes dans le CPNPC avancé [64].

Les DEG communs étaient principalement enrichies dans les voies de signalisation suivantes : *Cell cycle, DNA replication, Drug metabolism - cytochrome P450, Cytokine-cytokine receptor interaction*. Le cycle cellulaire est constitué d'une série complexe d'événements qui permettent à la cellule de croître et de proliférer [65]. La dérégulation du cycle cellulaire est une condition préalable à la formation de la plupart, voire de toutes les tumeurs malignes [66]. Les gènes impliqués dans la réplication de l'ADN (*DNA replication*) ont été récemment étudiés en tant que marqueurs prédictifs de la réponse à la chimiothérapie dans le CBNPC [67].

Par ailleurs Gharavi et al. ont prouvé que Les enzymes du cytochrome P450 (CYP450)

jouent un rôle clé dans le métabolisme de plusieurs aspects du cancer en raison de leur large spécificité de substrat, les CYP450 sont également importants dans le métabolisme des médicaments anticancéreux, car ils peuvent augmenter ou diminuer l'efficacité des médicaments en fonction de leur efficacité ou de leurs métabolites. Des études ont été menées sur la présence de différentes formes de CYP450 dans les tumeurs pulmonaires afin de déterminer si le métabolisme intra-tumoral des agents anticancéreux par le CYP450 peut avoir une influence sur la réponse de la tumeur à ces agents [68]. En supplément, Marrugal *et al.* ont exposé que les cytokines sont des protéines impliquées dans les processus inflammatoires et ont été proposées comme biomarqueurs de cancer du poumon et cibles thérapeutiques. En effet, il a été rapporté que certaines cytokines jouent un rôle important dans le développement, l'invasion et les métastases tumorales, telles que la cytokine IL-32. Cette dernière a été récemment proposée comme biomarqueur pronostique de l'adénocarcinome pulmonaire, car sa surexpression dans le tissu tumoral est corrélée à un plus grand nombre de métastases ganglionnaires [69].

Ensuite, à travers cette étude, nous avons trouvé des gènes clés (BUB1, ASPM, KIFF11, AURKB, CCNA2, NDC80, CCNB1, CCNB2) étaient significativement exprimés de manière différentielle entre les tissus tumoraux et les tissus normaux. L'expression de ces gènes peut indiquer que les chances de survie des patients atteints de cancer du poumon sont faibles.

L'ASPM (*abnormal spindle-like microcephaly-associated protein*) est une protéine associée aux microtubules du centrosome, qui joue un rôle dans plusieurs processus cellulaires ou développementaux, notamment dans la régulation de la neurogenèse et de la croissance du cerveau. L'ASPM peut maintenir des divisions symétriques des cellules neuroépithéliales. Une étude a constaté une expression étendue de l'ASPM dans plusieurs tissus tumoraux et son implication dans le développement et la progression de plusieurs cancers, notamment le cancer du poumon. Ya-Jing Yuan et *al.* ont montré que l'ASPM était impliquée dans le développement et la progression de l'adénocarcinome pulmonaire et était associée à un mauvais pronostic. En plus, la diminution d'ASPM inhibe considérablement la prolifération des cellules de *LSCC*, et ce gène pourrait représenter une nouvelle cible thérapeutique pour lutter contre le carcinome épidermoïde du poumon [70].

La protéine BUB1 (*BUB1 mitotic checkpoint serine/threonine kinase*) code pour une sérine/thréonine protéine kinase essentielle au processus de la mitose [71]. Le processus d'agrégation des chromosomes est contrôlé par BUB1, qui régule également le processus de localisation centromérique et est très important dans l'établissement et le maintien du processus d'attachement bipolaire efficace (pour réaliser l'attachement aux microtubules du fuseau) [57]. Elle est associée à un mauvais pronostic dans plusieurs types de cancer [71], y compris le cancer des poumons, car l'expression de l'ARNm de BUB1 était significativement élevée [57]. La surexpression de BUB1 dans l'adénocarcinome a révélé que BUB1 pourrait être un biomarqueur prometteur de *LUAD* et mérite une étude plus approfondie [71].

La protéine AURKB (*aurora kinase B*) est un membre de la sous-famille des kinases Aurora des sérine/thréonine kinases conservées avec une expression plus élevée dans les cellules tumorales que les cellules normales, et son surexpression a été associée aux caractéristiques biologiques et du diagnostic des cancers [72]. L'AURKB participe à la division cellulaire en régulant la condensation de la chromatine, la biorientation des chromosomes et la cytokinèse. En plus des fonctions de division cellulaire, une attention particulière a été accordée à l'AURKB dans différents contextes physiopathologiques. Il a été démontré que la sur expression d'AURKB est associée à une multiplication cellulaire infinie et à l'évitement du mécanisme apoptotique dans des multiples tumeurs. Yu et *al.* ont confirmé que les niveaux d'expression anormaux d'AURKB sont étroitement liés à une faible survie sans maladie et à la survie globale dans le CBNPC [73].

La CCNA2 (*Cell Cyclin A2*) appartient à la famille hautement conservée des cyclines. Elle est exprimée dans presque tous les tissus du corps humain. Plusieurs études ont montré que la CCNA2 pourrait augmenter le comportement agressif du cancer [74]. Dans le cas du cancer du poumon non à petites cellules où des niveaux d'expression élevés de la CCNA2 ont été proposés comme un biomarqueur de mauvais pronostic [75]. Zaho et al. ont montré que la CCNA2 est un biomarqueur potentiel de diagnostic et de pronostic pour le LUAD, et que son expression est positivement corrélée à l'efficacité de la thérapie immunitaire dans ce type de cancer [74].

La KIF11 (*kinesin family member 11*) appartient à la superfamille des kinésines et est impliquée dans la dynamique du fuseau. Elle code pour une protéine motrice moléculaire appelée Eg5, qui est impliquée dans la construction du fuseau bipolaire et la conduite de la mitose afin de favoriser la prolifération cellulaire. Il a été suggéré que la KIF11 est surexprimée dans plusieurs types de cancer et est considérée comme un biomarqueur de diagnostic et de pronostic potentiel en raison de son association significative avec la néoplasie. Une expression élevée de la KIF11 est corrélée à une faible survie globale chez les patients atteints d'adénocarcinome pulmonaire. Par conséquent, l'inhibition ciblée de la KIF11 est une stratégie thérapeutique potentielle pour les patients atteints de CBNPC [76].

La CCNB1 (*Cyclin B1*) appartient à une famille hautement conservée des cyclines [77]. Elle est surexprimée dans le CBNPC et sa forte expression est étroitement corrélée au ce type de tumeur, à la différenciation tumorale, à l'invasion vasculaire et au sexe masculin. Les patients atteints de CBNPC présentant une expression élevée de CCNB1 ont une survie globale plus courte que les patients présentant une faible expression de CCNB1. Notamment, une analyse multivariée a identifié l'expression de CCNB1 comme un facteur pronostique pour les patients atteints de CBNPC [78].

La CCNB2 (*Cyclin B2*) est un membre des protéines de la famille des cyclines, qui régulent les activités des kinases dépendantes des cyclines (CDK) et de différentes cyclines fonctionnent spatialement et temporellement dans des phases spécifiques du cycle cellulaire. La CCNB2 déclenche généralement la progression de la transition G2/M en activant la kinase CDK1 [79]. Gong et al. ont démontré que la surexpression de la protéine CCNB2 dans les poumons était liée à la progression clinique et au mauvais pronostic chez les patients atteints de CBNPC, suggérant que la CCNB2 pourrait être un biomarqueur du CBNPC [59].

La NDC80 (*kinetochore complex component*) est l'une des protéines du kinétochore externe [80]. Elle est située dans le Chr18p11.32 [81]. La NDC80 joue un rôle important dans la fixation stable des microtubules aux kinétochores [82]. Wang et al. ont démontré que ce gènes occupait une position centrale dans les tissus du CBPC. L'activation anormale de la NDC80 pourrait favoriser l'initiation du CBPC. Il pourrait constituer un biomarqueur pour le diagnostic précoce du CBPC. Par conséquent, NDC80 pourrait être considéré comme une cible thérapeutique pour le CBP [81].

Conclusion

Conclusion

L'objectif de cette étude était d'utiliser des méthodes bioinformatiques pour identifier les gènes essentiels liés au développement et à la progression du cancer du poumon. Les résultats obtenus à partir de cette étude sont les suivants :

- ✓ Identification de 179 gènes communs aux trois profils d'expression des patients, dont 78 gènes surexprimés et 101 gènes sous-exprimés.
- ✓ Les analyses de l'enrichissement fonctionnel ont révélé que les gènes identifiés étaient impliqués dans de nombreux processus biologiques, cellulaires et moléculaires associés au cancer des poumons.
- ✓ Huit gènes clés ont été identifiés (ASPM, BUB1, AURKb, CCNA2, KIF11, CCNB1, CCNB2 et NDC80), et ont montré une expression différentielle entre les tissus cancéreux et les tissus normaux.
- ✓ L'expression de ces huit gènes est associée à une mauvaise survie chez les patients atteints de cancer du poumon.

Au final, les résultats suggèrent que ces huit gènes clés identifiés pourraient être des biomarqueurs potentiels pour le diagnostic du cancer du poumon, ainsi pourraient être des cibles thérapeutiques pour le traitement du cancer des poumons.

Glossaire

Adénocarcinome : est un type de cancer qui se développe à partir de cellules glandulaires qui produisent du mucus ou d'autres substances liquides. Les glandes peuvent se trouver dans différents organes du corps, tels que les poumons, le pancréas, la prostate, le côlon, l'estomac et les seins.

BRAF1 : est un gène qui code pour la protéine BRAF (*B-Raf proto-oncogene, serine/threonine kinase*). Cette protéine est une kinase qui joue un rôle important dans la signalisation intracellulaire et la régulation de la croissance et de la division cellulaire.

Cytoscape : est un logiciel open-source et gratuit de visualisation, d'analyse et de modélisation de réseaux complexes. Il est principalement utilisé en bioinformatique pour la représentation graphique et l'analyse de données biologiques sous forme de réseaux tels que les réseaux de régulation génétique, les réseaux d'interactions protéine-protéine, etc...

DAVID : (*Database for Annotation, Visualization and Integrated Discovery*) est une plateforme de bioinformatique en ligne qui permet l'analyse de données biologiques à grande échelle. Elle fournit des outils pour l'analyse fonctionnelle des gènes, la classification des gènes en groupes fonctionnels et l'identification des voies de signalisation biologique associées.

EGFR : (*Epidermal Growth Factor Receptor*) : est un récepteur présent à la surface de certaines cellules qui régule leur croissance, leur division et leur survie. Il est activé par des molécules telles que le facteur de croissance épidermique et est impliqué dans le développement de certains cancers notamment les cancers du poumon, ce qui en fait une cible thérapeutique importante.

FC (fold change) : Le *fold-change* est le rapport du niveau moyen d'expression d'un gène dans une condition par rapport à une autre. Il est généralement exprimé en log (logarithme en base 2) afin de rendre symétriques les rapports par rapport à 1.

GO (Gene Ontology) : est une initiative majeure de bioinformatique visant à unifier la représentation des attributs des gènes et des produits géniques à travers toutes les espèces.

GEO (Gene Expression Omnibus) : est une base de données pour le profilage de l'expression génique et le profilage de la méthylation de l'ARN gérée par le *National Center for Biotechnology Information*.

GEPIA2 (*Gene Expression Profiling Interactive Analysis 2*) : est une plateforme en ligne permettant d'analyser les données d'expression génique de différents types de tumeurs et de tissus normaux. Elle permet aux utilisateurs de visualiser les niveaux d'expression de différents gènes dans les tissus cancéreux et normaux.

Hazard Ratio (HR): également connu sous le nom de rapport de risque instantané, est une mesure statistique qui compare le risque de survenue d'une maladie (ou autre événement) entre deux groupes différents.

KRAS (*kirsten Ras*) : est un gène humain qui code pour la protéine KRAS, qui fait partie de la famille des protéines Ras impliquées dans la régulation de la croissance, de la division et de la différenciation cellulaire. Les mutations du gène KRAS sont fréquemment observées dans différents types de cancer, ce qui en fait une cible thérapeutique importante.

KEGG (*Kyoto Encyclopedia of Genes and Genomes*) : est une base de données en ligne qui fournit des informations sur les gènes, les voies métaboliques et les maladies associées chez différents organismes. Elle est largement utilisée pour comprendre les interactions cellulaires et les réseaux de régulation moléculaire, et comprend des outils d'analyse bioinformatique pour aider à l'interprétation des données.

Kaplan-Meier Plotter : le diagramme de Kaplan-Meier Plotter, également connu sous le nom de courbe de survie, est un outil graphique utilisé pour représenter la probabilité de survie d'un groupe de sujets ou de patients dans une étude de suivi.

NCBI (*National Center for Biotechnology Information*) : est un centre national américain qui collecte, organise et diffuse des informations sur la biologie moléculaire, la génomique, la bio-informatique et les sciences connexes à travers plusieurs bases de données en ligne, telles que PubMed , GenBank, et BLAST .

P value : Une valeur p, qui signifie valeur de probabilité, est une mesure statistique comprise entre 0 et 1. Elle est utilisée pour un test d'hypothèse.

VEGF (*Vascular Endothelial Growth Factor*) : est une protéine impliquée dans la croissance et la régénération des vaisseaux sanguins. Il peut jouer un rôle dans le développement de maladies comme le cancer en favorisant la croissance de nouveaux vaisseaux sanguins nécessaires à la vascularisation de la tumeur.

Venn diagram : est un outil de représentation visuelle utilisé pour illustrer les relations entre différents ensembles. Popularisé par John Venn dans les années 1880.

Volcano plot : est un type de graphique utilisé en analyse de données pour visualiser les différences significatives entre deux groupes ou conditions. Les données sont représentées sous forme de points, où chaque point représente une variable

Références bibliographiques

1. Wang, K., et al., *Identification of differentially expressed genes in non-small cell lung cancer*. Aging, 2019. **11**(23): p. 11170-11185.
2. lalande, g.e., *cancers bronchopulmonaires mise au point* john libbey eurotext ed2006.
3. Frusch, N., L. Bosquee, and R. Louis, *Le cancer du poumon. Epidemiologie et facteurs etiologiques*. Revue Médicale de Liège, 2007. **62**(9).
4. Garnouat Merzak, K.S., *cancer broncho-pulmonaire primitif*. office des publications universitaires ed11-2016, place centrale ben aknoun-alger.
5. Bao, M. and G. Jiang, *Differential expression and functional analysis of lung cancer gene expression datasets: A systems biology perspective*. Oncol Lett, 2019. **18**(1): p. 776-782.
6. Kirienko, M., et al., *Radiomics and gene expression profile to characterise the disease and predict outcome in patients with lung cancer*. Eur J Nucl Med Mol Imaging, 2021. **48**(11): p. 3643-3655.
7. Tortora, G.J. and B. Derrickson, *Manuel d'anatomie et de physiologie humaines*2017: De Boeck supérieur.
8. Gautier-Isola, M., *Functional characterization of long non coding RNA induced by hypoxia and associated with non small cell lung carcinoma aggressiveness Caractérisation fonctionnelle de longs ARNs non codants induits par l'hypoxie et impliqués dans l'agressivité des cancers pulmonaires non à petites cellules*, 2020, Université Côte d'Azur.
9. Laurent Alleaume, I.D., Dr Pascale Dubray-Longeras, Marie-Agnès Faye., *les traitemets des cancers du poumon* in *cancer infonovember* 2017. p. 8.
10. Boulet, L.-P., et al., *Physiopathologie respiratoire appliquée*, 2014, Presses de l'Université Laval: [Québec, Québec].
11. Colonna, M., *Epidémiologie du cancer du poumon en France: incidence, mortalité et survie (tendance et situation actuelle)*. Revue des Maladies Respiratoires Actualités, 2016. **8**(5): p. 308-318.
12. Marouani.Abdellouche, M.K., L. Besbes, *Le cancer broncho-pulmonaire primitif*. Elsevier Masson SAS., 2015.
13. Didkowska, J., et al., *Lung cancer epidemiology: contemporary and future challenges worldwide*. Annals of Translational Medicine, 2016. **4**(8): p. 150.
14. Thandra, K.C., et al., *Epidemiology of lung cancer*. Contemp Oncol, 2021. **25**(1): p. 45-52.
15. Badheeb, A.M., et al., *2022 Book Cancer In The Arab World*2022.
16. Ferlay J, E.M., Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer. (2020) Available from: <https://gco.iarc.fr/today/data/factsheets/populations/12-algeria-fact-sheets.pdf>.
17. Wistuba, I.I. and A.F. Gazdar, *Lung cancer preneoplasia*. Annu. Rev. Pathol. Mech. Dis., 2006. **1**: p. 331-348.
18. Rudin, C.M., et al., *Small-cell lung cancer*. Nature Reviews Disease Primers, 2021. **7**(1): p. 3.

19. Byers, L.A. and C.M. Rudin, *Small cell lung cancer: Where do we go from here?* Cancer, 2015. **121**(5): p. 664-672.
20. Bernhardt, E.B. and S.I. Jalal, *Small Cell Lung Cancer*, in *Lung Cancer: Treatment and Research*, K.L. Reckamp, Editor 2016, Springer International Publishing: Cham. p. 301-322.
21. Jean-CHARLES Soria, s.V., Christophe Massard, Olivier Mir, *cours de chimiothérapie antitumorales et traitement médical du cancer*. John Libbey Eurotext ed2011.
22. Martin, F., *Le cancer Broncho-Pulmonaire*, in *Département de LA REUNION caractéristique clinique et démographique* 2016 Université de Bordeaux.
23. Hélène BOYLE, S.A.-L., Aude DURET , Bertrand FLEURY , Céline GUICHON, *Cancérologie et Radiothérapie*. ellipses edition marketing SA ed2011.
24. Travis, W.D., *Pathology of Lung Cancer*. Clinics in chest medicine, 2011. **32**(4): p. 669-692.
25. Vignaud, J.M., [*Squamous cell carcinoma, basaloid squamous cell carcinoma and adenosquamous carcinoma in the lung*]. Ann Pathol, 2016. **36**(1): p. 15-23.
26. Copin, M.C., [*Large cell carcinoma, lymphoepithelioma-like carcinoma, NUT carcinoma*]. Ann Pathol, 2016. **36**(1): p. 24-33.
27. Lantuejoul, S., *Classification anatomopathologique des cancers pulmonaires et de leurs préneoplasies*. Revue des Maladies Respiratoires Actualités, 2014. **6**(4): p. 346-357.
28. Py, C., *Revue Médicale Suisse : Syndromes paranéoplasiques : multiples facettes*. Revue Médicale Suisse, 2018. **14**(607): p. 1015-1020.
29. Morère, J.-F.o. and T. Le Chevalier, *Les cancers intrathoraciques*. Oncologie pratique / Jean-François Morère 2004, Paris: Springer.
30. amina, O., *etude physiopathologique du cancer du poumon chez les patients algériens a différents stades de la pathologie*, 2019, université des sciences et de la technologie houari boumadiene.
31. Masson, G.L.A.I.S.a.P.L., *Cecil Medicine Cancérologie* 2013.
32. Marina, d., *qualite de l'air*. 17 ed2013. 206.
33. Melloni, B., et al., *Radon et exposition domestique*. Rev Mal Respir, 2000. **17**: p. 1061-1071.
34. Clin, B. and J.C. Pairon, *Facteurs de risques professionnels des cancers broncho-pulmonaires, données épidémiologiques, repérage des expositions et enjeux médico-sociaux : l'amiante comme paradigme*. Revue des Maladies Respiratoires Actualités, 2022. **14**(2, Supplement 1): p. 2S5-2S9.
35. MAZOUZI, C., et al., *Le cancer du poumon*. Prostate. **2**(578): p. 4-9.
36. Matrat, M., et al., *Facteurs de risque professionnels des cancers du poumon et des voies aéro-digestives supérieures: l'étude ICARE*.
37. Molina, J.R., et al., *Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship*. Mayo Clin Proc, 2008. **83**(5): p. 584-94.
38. Xue, Y., et al., *Air pollution: A culprit of lung cancer*. Journal of Hazardous Materials, 2022. **434**: p. 128937.

39. Chassagnon, G., S. Bennani, and M.P. Revel, *Nouvelle classification TNM des cancers du poumon non à petites cellules*. Revue de Pneumologie Clinique, 2017. **73**(1): p. 34-39.
40. Diane, m., *profils moléculaires des cancers bronchiques non a petits cellules a la reunion 2019*.
41. T.Berghmans, B.G., *revue des maladies respiratoires actualités*, e.m. sas, Editor 2020.
42. Goldstraw, P., et al., *The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer*. Journal of Thoracic Oncology, 2016.
43. Bennani Smires, C., *TEP-TDM en pratique*, 2022, Elsevier Masson: Issy-les-Moulineaux, France.
44. Laffon, E., H. De Clermont, and L. Bordenave, *TEP-TDM et thorax*. Revue des maladies respiratoires, 2010. **27**(10): p. 1275-1280.
45. duperray, i.p.e.m., *l'imagerie par résonance magnétique (IRM)*, 2020.
46. Ferretti, G., et al., *Imagerie radiologique et TEP Scanner dans les cancers du poumon*. Journal de Radiologie, 2008. **89**(3): p. 387-402.
47. Ihadadene, D., et al., *Rôle de l'endoscopie bronchique dans le diagnostic du cancer bronchique*. Revue des maladies respiratoires, 2016. **33**: p. A214.
48. Homasson, J., et al., *Endoscopie interventionnelle par bronchoscopie souple*. J Fran Viet Pneu, 2010. **1**: p. 23-8.
49. Norouzi, M. and P. Hardy, *Clinical applications of nanomedicines in lung cancer treatment*. Acta Biomaterialia, 2021. **121**: p. 134-142.
50. Grand, B., *La chirurgie du cancer du poumon : vidéochirurgie, robot, chirurgie infra-lobaire*. Revue des Maladies Respiratoires Actualités, 2013. **5**(5): p. 490-495.
51. Boisselier, P., et al., *Intérêt de la radiothérapie stéréotaxique pour les tumeurs primitives du poumon localisées de stade T1-T2*. Cancer/Radiothérapie, 2022. **26**(6): p. 755-759.
52. Faivre-Finn, C., *Cancers bronchiques non à petites cellules localement avancés : quelle radiothérapie en 2014 ?* Revue des Maladies Respiratoires Actualités, 2014. **6**(4): p. 421-430.
53. S. Vignot a, G.T.a., D. Solub a, C. Wilkowsky *Thérapies moléculaires ciblées* -. 20/08/2015.
54. Screening, P. and P.E. Board, *Lung Cancer Prevention (PDQ®)*, in *PDQ Cancer Information Summaries [Internet]*2022, National Cancer Institute (US).
55. Bade, B.C. and C.S. Dela Cruz, *Lung Cancer 2020: Epidemiology, Etiology, and Prevention*. Clin Chest Med, 2020. **41**(1): p. 1-24.
56. Akhtar, N. and J.G. Bansal, *Risk factors of Lung Cancer in nonsmoker*. Current Problems in Cancer, 2017. **41**(5): p. 328-339.
57. Li, X., et al., *Comprehensive analysis of BUBs gene family in lung adenocarcinoma with immunological analysis*. Aging, 2023. **15**(3): p. 810-829.
58. Lemjabbar-Alaoui, H., et al., *Lung cancer: Biology and treatment options*.

- Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2015. **1856**(2): p. 189-210.
59. Gong, K., et al., *Identification and integrate analysis of Key biomarkers for diagnosis and prognosis of non-small cell lung cancer based on bioinformatics analysis*. Technology in Cancer Research & Treatment, 2021. **20**: p. 15330338211060202.
60. Madelaine, J. and G. Zalzman, *Biologie des cancers bronchiques*. EMC - Pneumologie, 2005. **2**(1): p. 9-31.
61. Masuda, A. and T. Takahashi, *Chromosome instability in human lung cancers: possible underlying mechanisms and potential consequences in the pathogenesis*. Oncogene, 2002. **21**(45): p. 6884-6897.
62. Jefford, C.E. and I. Irminger-Finger, *Mechanisms of chromosome instability in cancers*. Critical reviews in oncology/hematology, 2006. **59**(1): p. 1-14.
63. Dumontet, C. and M.A. Jordan, *Microtubule-binding agents: a dynamic field of cancer therapeutics*. Nature Reviews Drug Discovery, 2010. **9**(10): p. 790-803.
64. Sève, P. and C. Dumontet, *β -tubuline de classe III et cancer du poumon non à petites cellules*. Revue des maladies respiratoires, 2010. **27**(4): p. 383-386.
65. Mirjolet, J.-F., *Cycle cellulaire et pharmacologie des anticancéreux*. Cycle cellulaire et cytométrie en flux, 2010: p. 151.
66. Müller-Tidow, C., et al., *Cyclin E is the only cyclin-dependent kinase 2-associated cyclin that predicts metastasis and survival in early stage non-small cell lung cancer*. Cancer research, 2001. **61**(2): p. 647-653.
67. Saviozzi, S., et al., *Non-small cell lung cancer exhibits transcript overexpression of genes associated with homologous recombination and DNA replication pathways*. Cancer research, 2009. **69**(8): p. 3390-3396.
68. Gharavi, N. and A.O. El-Kadi, *Expression of cytochrome P450 in lung tumor*. Current drug metabolism, 2004. **5**(2): p. 203-210.
69. Marrugal, Á., et al., *Proteomic-Based Approaches for the Study of Cytokines in Lung Cancer*. Disease Markers, 2016. **2016**: p. 2138627.
70. Yuan, Y.-J., et al., *Abnormal spindle-like microcephaly-associated protein (ASPM) contributes to the progression of Lung Squamous Cell Carcinoma (LSCC) by regulating CDK4*. Journal of Cancer, 2020. **11**(18): p. 5413.
71. Wang, L., et al., *Bioinformatics analysis of BUB1 expression and gene regulation network in lung adenocarcinoma*. Transl Cancer Res, 2020. **9**(8): p. 4820-4833.
72. Sang, L., et al., *Bioinformatics analysis of aberrantly methylated-differentially expressed genes and pathways in hepatocellular carcinoma*. World J Gastroenterol, 2018. **24**(24): p. 2605-2616.
73. Chen, B., et al., *Identification of prognostic markers by weighted gene co-expression network analysis in non-small cell lung cancer*. Bioengineered, 2021. **12**(1): p. 4924-4935.
74. Jiang, A., et al., *CCNA2 as an immunological biomarker encompassing tumor microenvironment and therapeutic response in multiple cancer types*.

- Oxidative Medicine and Cellular Longevity, 2022. **2022**.
75. Yang, X., et al., *Identification of a transcription factor- cyclin family genes network in lung adenocarcinoma through bioinformatics analysis and validation through RT- qPCR*. Experimental and Therapeutic Medicine, 2023. **25**(1): p. 1-14.
 76. Xue, T., et al., *Cyclovirobuxine D inhibits growth and progression of non- small cell lung cancer cells by suppressing the KIF11- CDC25C- CDK1- CyclinB1 G2/M phase transition regulatory network and the NFκB/JNK signaling pathway*. Int J Oncol, 2023.
 77. Ding, K., et al., *CCNB1 is a prognostic biomarker for ER+ breast cancer*. Medical Hypotheses, 2014. **83**(3): p. 359-364.
 78. Wang, S., et al., *MicroRNA- 718 serves a tumor- suppressive role in non- small cell lung cancer by directly targeting CCNB1 Retraction in /10.3892/ijmm.2021.5013*. Int J Mol Med, 2020. **45**(1): p. 33-44.
 79. Qian, X., et al., *CCNB2 overexpression is a poor prognostic biomarker in Chinese NSCLC patients*. Biomedicine & Pharmacotherapy, 2015. **74**: p. 222-227.
 80. Wei, R., et al., *Bioinformatic analysis revealing mitotic spindle assembly regulated NDC80 and MAD2L1 as prognostic biomarkers in non-small cell lung cancer development*. BMC Medical Genomics, 2020. **13**(1): p. 1-14.
 81. Wang, X., et al., *Monastrol Targeted KIF11 Showed Potential Treatment Effective of Small Cell Lung Cancer*. 2019.
 82. Kaneko, N., et al., *siRNA-mediated knockdown against CDCA1 and KNTC2, both frequently overexpressed in colorectal and gastric cancers, suppresses cell proliferation and induces apoptosis*. Biochemical and biophysical research communications, 2009. **390**(4): p. 1235-1240.